



Universiteit  
Leiden  
The Netherlands

## Structure and function of the UVDE repair protein

Paspaleva, K.

### Citation

Paspaleva, K. (2009, April 1). *Structure and function of the UVDE repair protein*. Retrieved from <https://hdl.handle.net/1887/13711>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/13711>

**Note:** To cite this publication please use the final published version (if applicable).

# Structure and function of the UVDE repair protein

Keti Paspaleva



# Structure and function of the UVDE repair protein

**PROEFSCHRIFT**

ter verkrijging van de graad van Doctor aan de Universiteit Leiden,  
op gezag van Rector Magnificus Prof. mr. P.F. van der Heijden,  
volgens besluit van het College voor Promoties  
te verdedigen op woensdag 1 april 2009  
klokke 16.15

door

**Keti Paspaleva**

geboren te Bourgas in 1978

## **PROMOTIECOMMISSIE**

Promotor: Prof. dr. J.P. Abrahams

Co-promotor: Dr. N. Goosen

Referent: Dr. H. Vrieling

Overige leden: Prof. dr. M. Noteborn

Prof. dr. J. Brouwer

Prof. dr. T. Sixma

Dr. E. Thomassen

Prof. dr. B. van Houten (University of Pittsburgh, USA)

For my parents and my beloved granny



## **CONTENTS**

<b>Chapter 1</b>	General introduction	9
<b>Chapter 2</b>	Crystal structure of the DNA repair enzyme UV damage endonuclease	45
<b>Chapter 3</b>	Involvement of a carboxylated lysine in UV damage endonuclease from <i>Thermus thermophilus</i>	61
<b>Chapter 4</b>	Damage recognition by UV damage endonuclease from <i>Schizosaccharomyces pombe</i>	79
<b>Chapter 5</b>	Active site organisation of UVDE - a Mn <sup>2+</sup> dependent nuclease	103
	<b>Summary and general discussion</b>	129
	<b>Samenvatting en algemene discussie</b>	132
	<b>Curriculum vitae</b>	135
	<b>List of publications</b>	135
	<b>Acknowledgements</b>	136



# General introduction Damage recognition in DNA repair systems

---

## 1. INTRODUCTION

DNA carries genetic information essential for all processes of life. The DNA in all living organisms is under a constant attack by numerous endogenous and environmental agents.

### 1.1 Endogenous DNA damaging sources

At physiological conditions DNA continuously interacts with oxygen and water, which can lead to creation of DNA lesions. Major sites of oxidative and hydrolytic damage are the amino groups of cytosine, adenine and guanine. The loss of the amino groups (deamination) occurs spontaneously and converts affected DNA bases into uracil, hypoxanthine and xanthine, respectively. Another target for the hydrolytic and oxydative sources is the N-glycosylic bond. Cleavage of this bond results in a base loss and creation of an apurinic/aprimidinic (AP) site in the DNA. The pyrimidine nucleotides are considerably more stable than the purine ones, since cytosines and thymines are lost at a much lower rate, compared to the adenines and the guanines (Lindahl and Nyberg, 1974). Other well known DNA damaging agents are reactive oxygen species (ROS), which include oxygen ions, free radicals and peroxides. ROS can be formed as a by-product of the cellular metabolism or induced by exogenous sources such as ionizing radiation. One of the most potent ROS is the hydroxyl radical ( $\bullet\text{OH}$ ). It targets the double bonds of DNA bases as well as the deoxyribose sugars, from which it abstracts hydrogen atoms. The  $\bullet\text{OH}$  attack on the double bond of thymine bases yields a 6-hydroxythymine radical, which can react with  $\text{O}_2$ , forming a thymine glycol (Dempfle and Linn, 1982). The  $\bullet\text{OH}$  interactions with the DNA sugar residues may result in DNA fragmentation, base loss or strand breaks.

## 1.2 Exogenous DNA damaging sources

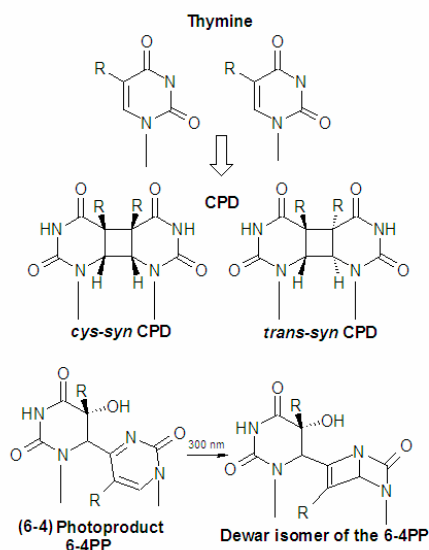
Exogenous (environmental) sources include ionizing radiation, cross-linking agents, alkylating agents and aromatic compounds. One of the most common natural sources for induction of DNA lesions is the UV component of sunlight.

### UV light

Depending on its wavelength, UV irradiation has been divided into: UVA (320 to 400 nm), UVB (295 to 320 nm) and UVC (100 to 295 nm). Solar UV light, which reaches Earth, consists mainly of UVA and UVB, since UVC is efficiently blocked by the ozone layer.

The ultraviolet light can be absorbed by two adjacent pyrimidines in the DNA (Yoon *et al.*, 2000), creating covalent bonds between them. Two common UV products are the cyclobutane pyrimidine dimer (CPD) and the 6-4 photoproduct (Figure 1).

The cyclobutane pyrimidine dimer is a four membered ring structure, formed between the C-5 and C-6 atoms of two neighbouring pyrimidines. T-T dimers (thymine dimers) are the most abundant of CPDs, although C-T, T-C and C-C dimers are also possible (Mitchell *et al.*, 1992). Twelve isomeric forms of T-T CPDs exist, however, only two with the conformations *cis-syn* and *trans-syn* occur with significant yields (David, 2001). The 6-4 Photoproduct (6-4PP) or also designated as pyrimidine-pyrimidone (6-4) photoproduct (Figure 1) is formed between C-6 of the 5' pyrimidine and C-4 of the 3' pyrimidine. Irradiation of the 6-4PP with 300-350 nm light converts it to a Dewar isomer (Figure 1), lesion that might be of a considerable biological relevance, since solar UV-B is in that wavelength range (Mitchel and Narin, 1989), (Taylor *et al.*, 1988).



**Figure 1.** Structure of *cis-syn* and *trans-syn* cyclobutane pyrimidine dimer (CPD), (6-4) pyrimidone photo-product (6-4PP) and its Dewar isomer. R represents a  $-CH_3$  group.

It was shown (Yoon *et al.*, 2000) that the frequency of occurrence of the CPD dimer is much higher (70 % - 80 %) than the induction of the 6-4PP (20 % - 30 %). However, the 6-4PP was revealed to be extremely mutagenic in contrast to the CPD (Mitchel and Narin, 1989). The CPD and the 6-4PP lesions are produced efficiently by UVC and UVB irradiation of DNA. The absorption of UVA photons by DNA is rather weak and it is thought that indirect mechanisms are responsible for its biological effects. The DNA damage induced by UVA may involve endogenous chromophores serving as radiation-absorbing intermediates. These can generate reactive oxygen species (ROS), which may further damage DNA (Friedberg *et al.*, 2005).

The CPD and the 6-4PP exhibit significant differences in the DNA structural distortion. For the CPD different degree of helix deformation has been reported in literature. The bending induced by this lesion was initially reported as 30° by a circularization assay (Husain *et al.*, 1988), but later as 7° in a phased multimer gel electrophoretic assay (Wang and Taylor, 1991). The CPD-containing DNA duplexes have also been a subject of theoretical studies (Pearlman *et al.*, 1985), which predicted that CPD might cause bending from 6° up to 28° (Liu *et al.*, 2000). More recently, the crystal structure of a DNA decamer containing a *cis-syn* thymine dimer showed 30° bending and 9° unwinding (Park *et al.*, 2000). Incorporation of a *cis-syn* CPD into double-stranded DNA was also shown to destabilise the duplex by 1.5 kcal/mol (Jing *et al.*, 1998). Although a reasonable Watson and Crick base pairing can still occur at the 3'T of the TT dimer, the base pairing of the 5'T is severely weakened (Park *et al.*, 2000).

The 6-4PP causes 44° bending and a much higher degree of distortion compared to the CPD. The DNA duplex is destabilised by 6 kcal/mol and although initially it was reported that both pyrimidines of the 6-4PP lose their ability to form hydrogen bonds with the opposite strand (Mitchel and Narin, 1989), later NMR studies suggested that this happens only at the 3'-side of the (6-4) lesion (Kim and Choi, 1995).

The Dewar isomer induces a substantially different change in the overall DNA structure compared to the 6-4PP: for example an overall helical bending of 21° rather than the 44° caused by the 6-4PP (Lee *et al.*, 1998).

DNA damages interfere with important processes like transcription and replication. Organisms cannot tolerate such genome threats and a variety of repair strategies have evolved to remove UV- and other DNA lesions (Wilson and Thomson, 1997). Four important pathways for repair of damaged bases in the DNA are: direct reversal, base excision repair (BER), nucleotide excision repair (NER) and UV damage endonuclease (UVDE) repair.

## 2. DIRECT REVERSAL

An example of simple and effective type of DNA repair is direct reversal. It involves damage specific enzymes, which restore the DNA to its native state in a single-step reaction. Such a simple pathway is kinetically faster than the multistep reaction, catalysed by multiprotein complexes

and is considered to be essentially error free (Friedberg *et al.*, 2005). Here two examples of direct reversal will be discussed: the enzymatic removal of alkyl groups from the DNA by methyl transferases and the splitting of the UV induced CPD and 6-4PP by photolyases.

## 2.1 Methyl-transferases

The DNA repair protein O<sup>6</sup>-alkylguanine-DNA alkyltransferase (AT) plays an important role in the cellular defence against alkylating agents. Alkyl adducts at the O<sup>6</sup>-position of guanine are one of the most mutagenic lesions, although they are a relatively minor product in relation to other sites of DNA alkylation (Sedgwick *et al.*, 2007), (Jackson *et al.*, 1997).

Alkyltransferases (AT) act to remove alkyl groups from the O<sup>6</sup>-position of a guanine through irreversible, single-step transfer of adducts to an active site cysteine residue (Samson, 1992), (Demple *et al.*, 1985). In addition, AT can also remove the alkyl group from the O<sup>4</sup>-position of an O<sup>4</sup>-methylthymine (Paalman *et al.*, 1997). The O<sup>4</sup>-methylthymine is a rather rare methylation product, which can form an incorrect pair with guanine, resulting in a TA to CG transition (Sedgwick *et al.*, 2007). Upon alkylation of the active site cysteine, the AT enzymes become highly susceptible to proteolysis (Kanugula *et al.*, 1998) and therefore are called suicide proteins.

The O<sup>6</sup>-alkylguanine-DNA alkyltransferase enzymes can be found in both eukaryotic and prokaryotic organisms and all proteins exhibit the same active site sequence V(I)PCHRV(I) (Friedberg *et al.*, 2005). The best characterised AT proteins are those from *Escherichia coli* and man. *E. coli* contains two alkyl transferases: the product of the *ogt* gene and the product of the *ada* gene. The Ogt protein is constitutively expressed and preferentially removes bulky alkyl adducts such as O<sup>6</sup>-benzylguanine (Goodtzova *et al.*, 1997). In contrast, the *ada* gene is activated upon exposure to alkylating agents and the Ada enzyme is specialised in the repair of O<sup>6</sup>-methylguanine. The crystal structure of the 19 kDa C-terminal domain of the *E. coli* protein (Ada-C) has been determined (Moore *et al.*, 1994). This 178 amino acids C-terminal fragment is thought to be responsible for the methyltransferase activity on O<sup>6</sup>-methyl-guanine-DNA. It houses a guanine specific binding pocket, in which, as predicted by DNA modelling, the methylated nucleotide can only bind if it is flipped out of the helix (Moore *et al.*, 1994). So far the only co-crystal structure of an O<sup>6</sup>-alkylguanine-DNA alkyltransferase bound to its substrate is available for the human homologue. The 21 kDa human protein (hAT) is homologous to the C-terminal domain of the Ada protein from *E. coli* (Ada-C). The structure of the hAT enzyme bound to its substrate (Daniels *et al.*, 2004) elucidates the mechanism of damage verification and activity of this class of enzymes. Binding of the protein recognition helix widens the DNA minor groove and Arg128, situated at its beginning, promotes flipping of the target nucleotide out of the base stack into the hAT active site (Figure 2).

The arginine side chain stacks between the bases that flank the substrate nucleotide and can form a hydrogen bond with the orphaned cytosine, thus largely compensating for the loss of the nucleotide from the base stack. This 'arginine finger' stabilises the extrahelical DNA conformation and may also actively push nucleotides from the base stack during the DNA scanning (Daniels *et al.*, 2004).



**Figure 2.** Crystal structure of the human O<sup>6</sup>-alkylguanine-DNA alkyltransferase (hAT) bound to an O<sup>6</sup>-methylguanine.

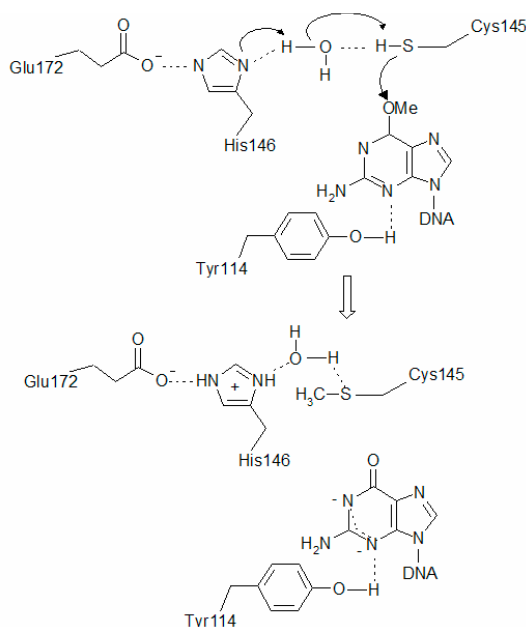
The extrahelical O<sup>6</sup>-methylguanine and the Arg finger are indicated in blue.

Upon flipping, the extrahelical base is accommodated in a hydrophobic pocket, which provides geometric exclusion for all DNA bases, except guanine. Furthermore, hydrogen bonds are formed between the carbonyl groups of Cys145, Val148 and the amino group of a normal or methylated guanine, thus providing the enzyme selectivity for these bases. From the crystal structure, however, it is not entirely clear how hAT discriminates between the unmodified and the methylated guanine. It is currently thought (Daniels *et al.*, 2004) that the difference in affinity for O<sup>6</sup>-methylguanine over guanine is based solely on the larger hydrophobic surface derived from alkylation.

The next step in the hAT repair mechanism is the deprotonation of the active site cysteine. In this process (Figure 3), His146 acts as water-mediated general base to deprotonate Cys145, which serves as nucleophile in the dealkylation reaction. Donation of a hydrogen bond from Tyr114 to N3 of the target guanine may also promote the reaction by reducing the negative charge on the methylated guanine (Figure 3).

The crystal structure of the hAT enzyme reveals the presence of one Zn<sup>2+</sup> ion. The metal ion lies in close proximity to the active site but is ~ 20 Å away from the reactive cysteine. The role of the Zn<sup>2+</sup> is proposed to be structural, since it stabilises the interface between the N- and the C-terminal domains of the protein (Daniels *et al.*, 2004).

In conclusion, the hAT enzyme does not utilize a metal cofactor for the deprotonation of the active site cysteine. Instead, as discussed, protein side chain (His146) acts as a water-mediated base and assists the deprotonation of Cys145.



**Figure 3.** Reaction mechanism of the human AT alkyltransferase.

## 2.2 Photoreactivation

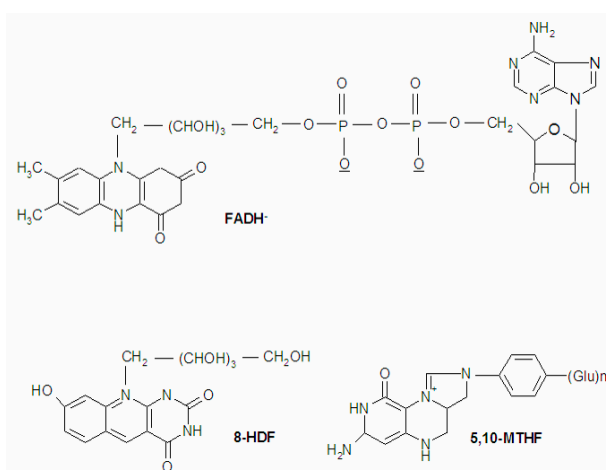
Photoreactivation is an efficient light-dependent process, which uses UVA (320 - 400 nm) and blue light (400 - 500 nm) to monomerise CPD dimers and 6-4 photoproducts. Photolyases are widely spread in nature and are monomeric proteins with molecular masses in the 53 - 66 kDa range, depending on the organism (Weber, 2005). Photolyases are distinguished by their different substrate specificity: CPD photolyase binds and repairs only CPD lesions, while the (6-4) photolyase reverses only the 6-4PP (Todo, 1999). All known photolyases contain non-covalently bound flavin adenine dinucleotide (FAD) as redox-active cofactor and an antenna pigment, which may differ (Zhao *et al.*, 1997).

Photolyases can be found in various organisms: bacteria, yeast, insects and many vertebrates, including aplacental mammals. Some species like *Drosophila melanogaster* contain both a CPD and a (6-4) photolyase, however, most organisms possess only one type of photolyase (Goosen and Moolenaar, 2007). Although all placental mammals do not have any photolyase activity, two genes with a high similarity to the *Drosophila* (6-4) photolyase have been identified in the human genome and designated as hCRY1 and hCRY2 (Kobayashi *et al.*, 1998). They do not have a DNA repair function, but instead, act as photoreceptors of the circadian clock. CRY genes were initially identified in the plant *Arabidopsis thaliana* and the gene products named CRY1 and CRY2 from 'cryptochrome'. CRY1 is involved in the plant elongation, while CRY2 regulates the flowering in response to blue light (Guo *et al.*, 1998).

### CPD photolyases

CPD photolyases are the best studied type of DNA photolyases and have been divided into two classes (I and II), based on their amino acid sequence similarity. Class I photolyases are found in many microorganisms, while most of class II photolyases are found in higher eukaryotes (Yasui *et al.*, 2001). Recently, a CPD photolyase specific for CPD in ssDNA has been reported in *Vibrio cholerae* and designated as Cry1. Cry1 is a member of the Cry-DASH subfamily of cryptochromes and contains MTHF and FADH as chromophores (Worthington *et al.*, 2003).

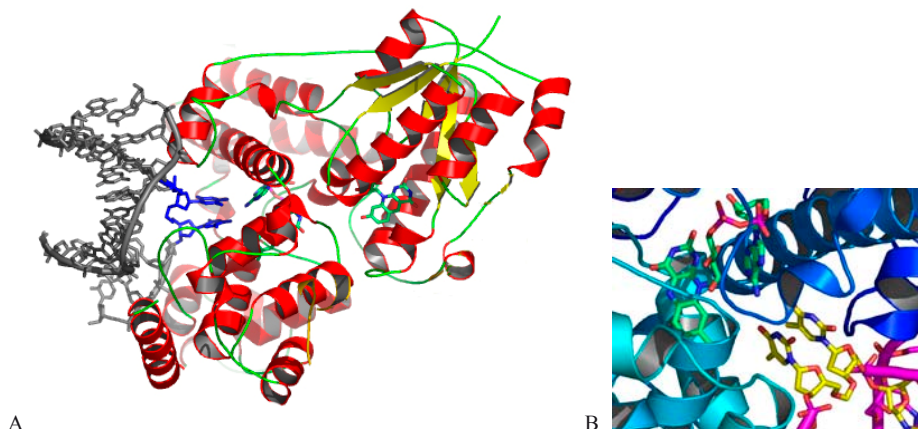
CPD photolyases generally have two kinds of chromophores. One is catalytic cofactor (FADH<sub>2</sub>), which directly interacts with the CPD substrate in a photo-repair reaction. The other is light-harvesting cofactor, which acts as an antenna to harvest light, transferring the energy to the catalytic cofactor (Kim *et al.*, 1992 and 2001). Class I photolyases are categorised according to their second chromophore into either a deazaflavin- or a folate-type. A deazaflavin-type photolyase has an 8-hydroxy-5-deazaflavin (8-HDF) as light-harvesting cofactor, while the folate-type photolyase has 5, 10-methenyltetrahydrofolic acid (MTHF) (Figure 4).



**Figure 4.** Structures of chromophores found in the pyrimidine dimer-DNA photolyases.

The folate class of DNA photolyases contains FADH<sup>-</sup> and 5,10-MTHF. The deazaflavin class of DNA photolyases contains FADH<sup>-</sup> and 8-HDF.

The recognition of CPD lesions by the DNA photolyases is elucidated by the crystal structure of the protein from *Anacystis nidulans* (Figure 5) in a complex with 14-nucleotide oligomer DNA duplex with a CPD in a central position (Mees *et al.*, 2004). The structure is solved with 1.8 Å resolution. Despite the extensive structural distortion of the DNA upon photolyase binding, the protein itself undergoes only minor changes. The CPD lesion is flipped (Figure 5A) and inserted into the active site of the enzyme in such a way that the thymine dimer is suitably positioned to form hydrogen bonds with the catalytic FADH<sup>-</sup> (Figure 5B).



**Figure 5.** The *A. nidulans* CPD photolyase in a complex with CPD containing DNA substrate.

A. The *A. nidulans* CPD photolyase-DNA complex. The CPD dimer (blue) is flipped out of the DNA helix.

B. Zoom in the enzyme active site. The flipped CPD lesion (yellow) is situated in close proximity to the FADH $\cdot$  (colored green).

Salt bridges and hydrogen bonds are extensively formed along the protein surface and the phosphates of the DNA substrate. The crucial role of these interactions for the formation of the photolyase-CPD DNA complex is underlined by a mutation of a conserved arginine (*A. nidulans* R350 and *E. coli* R342), which forms direct and water-mediated hydrogen bonds with the O<sup>2</sup> and O<sup>4</sup> oxygens of the thymine bases (Essen *et al.*, 2006). The arginine to alanine substitution is found to exhibit a severe phenotype causing a 32-fold decrease in the protein-DNA complexes and drop in activity from 98 % to about 60 %.

In the co-crystal structure of the CPD photolyase from *A. nidulans* the DNA is bend to about 50° and partially unwound. Residues Gly397 to Phe406 of the protein occupy the vacant space. The adenines complementary to the CPD are distorted, but still in an intrahelical position (Mees *et al.*, 2004).

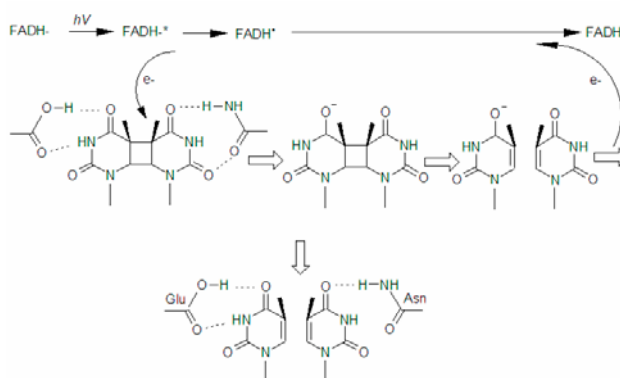
In summary, upon damage recognition the CPD photolyase flips the CPD lesion into its active site, where the DNA repair occurs by light-driven transfer of an electron from the excited FADH $\cdot^*$  to the CPD lesion (Figure 6). About 125 kJ/mol of the 240 kJ/mol of energy that is captured upon photon absorption is consumed during this initial electron transfer step (Carell *et al.*, 2001). After electron capture, the splitting of the CPD lesion proceeds rapidly within 0.6 ns (Sancar, 2000). After DNA repair, the thymine pair has to flip back into the duplex DNA to form hydrogen bonds with the complementary adenines. This relaxation of the DNA backbone proceeds at a much lower speed than the repair of the CPD lesion itself, as was shown by spectroscopy methods (Essen *et al.*, 2006).

*(6-4) photolyases*

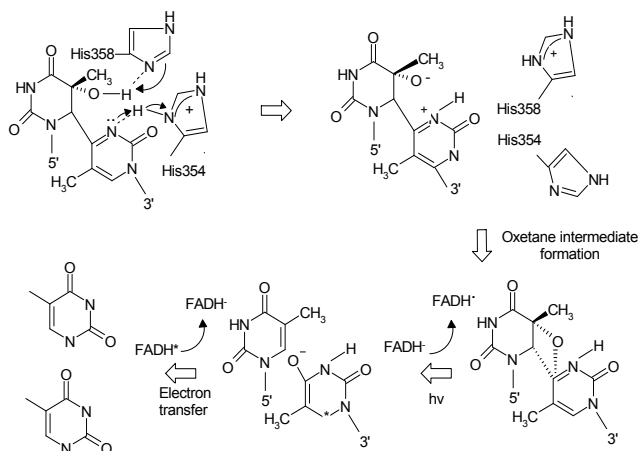
The fact that both the CPD and the 6-4PP photolyases use FAD for catalytic factor suggests that the basic repair mechanism of both groups might be similar. In both cases electron donation is needed in order to convert the TT dimers (CPD or 6-4PP) to its original form. For the 6-4PP, however, this process is more complicated. The 6-4 photoproduct is structurally different from the cyclobutane pyrimidine dimer (CPD), since it involves creation of a bond between the C-6 of the 5' base and the C-4 of the 3' base (Figure 1). If an enzyme would break the 6-4 C-C bond the bases would not be restored to their original forms.

A (6-4) photolyase was described for the first time in *Drosophila melanogaster* (Todo *et al.*, 1996), and later on, similar activity was observed in some vertebrates and plants. Currently, the cDNAs of (6-4) photolyase have been cloned from *D. melanogaster* (Todo *et al.*, 1996), *Xenopus laevis* (Todo *et al.*, 1997), *Arabidopsis thaliana* (Nakajima *et al.*, 1998) and *Danio rerio* (Zebra fish) (Kobayashi *et al.*, 2000). A crystal structure of a (6-4) photolyase with its 6-4PP substrate is not yet available.

Mutational analysis done on the *X. laevis* (6-4) photolyase (Hitomi *et al.*, 2001) outlined three residues that are likely to be involved in the enzyme catalytic function. His354, Leu355 and His358 are highly conserved and computer modelling of the enzyme active site predicted them to be in a close contact with the 6-4PP damage. The model used the available structural data from the CPD photolyases in order to predict the possible organisation of the (6-4) photolyase active site. Based on the computer modelling, a catalytic mechanism was proposed (Hitomi *et al.*, 2001) (Figure 7) in which His354 and His358 form hydrogen bonds with the N<sup>3</sup> atom of the 3'-pyrimidone and the hydroxyl group on the 5'-pyrimidine, respectively. While His358 abstracts a proton from the -OH group of the 5'-base, His354 protonates the N<sup>3</sup> of the 3'-base to generate a highly electrophilic iminium ion (Figure 7). Upon formation of the iminium ion, the reaction proceeds to formation of an oxetane intermediate. The intermediate is then converted to the native thymines by first absorption of excitation energy (from FADH) and later transfer of an electron back to the flavin radical (Hitomi *et al.*, 2001).



**Figure 6.** Reaction mechanism of the CPD photolyase from *A. nidulans*.



**Figure 7.** Proposed in literature mechanism for the 6-4PP photolyases activity.

### 3. NUCLEOPHILIC SUBSTITUTION: A COMMON MECHANISM FOR DNA HYDROLYSIS

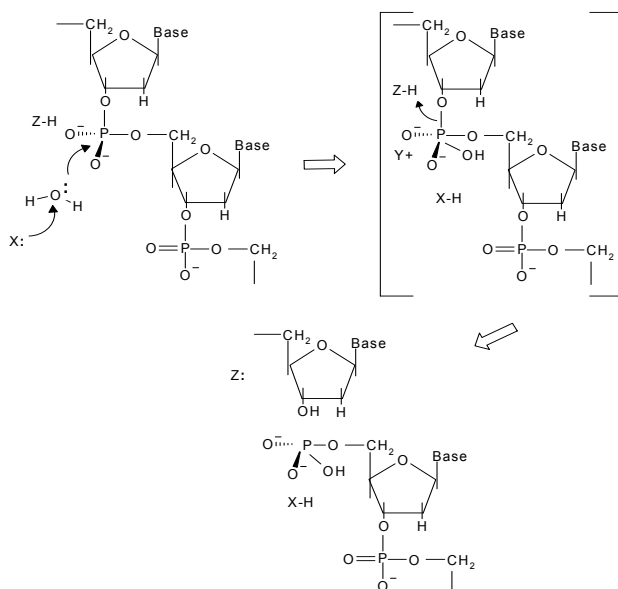
All direct reversal DNA repair processes have a common problem – they can only repair alterations of the DNA bases. What organisms need are more general mechanisms, capable of correcting numerous lesions by removing the damaged nucleotides from the DNA backbone. This requirement is met by the excision repair pathways (Base Excision Repair and Nucleotide Excision Repair). A general feature in all excision repair mechanisms is the removal of the nucleotides containing the damage by incision of the DNA backbone.

The mechanism of a DNA cleavage is referred to as Nucleophilic substitution and the general reaction can be given as:



First, the electron-rich nucleophile (OH<sup>-</sup>) attacks the DNA backbone and forms an intermediate product (OH-DNA-LG). In this intermediate LG designates the so called leaving group, which at the collapse of the unstable intermediate, departs with an electron pair (LG<sup>-</sup>). Since the addition of the nucleophile and the elimination of the leaving group take place simultaneously (Carey and Sundberg, 2000), the nucleophilic substitution of DNA is classified as an S<sub>N</sub>2 reaction (Jencks, 1981).

At a physiological pH, the DNA backbone has a large barrier for cleavage (Galburt and Stoddard, 2002), since it is negatively charged and electrostatically repels potential attacking nucleophiles (Westheimer, 1987). In order for the energy barrier to be overcome and the DNA to be cleaved, several elements are required. These include a nucleophilic group (OH<sup>-</sup>)



**Figure 8.** Schematic representation of the DNA cleavage performed by nucleases.

X represents a general base, Y represents a Lewis acid, Z-H represents a general acid, i and the brackets show the pentacoordinate intermediate, which is unstable.

to attack the DNA backbone, a basic moiety to activate and position the nucleophile (X in the scheme, Figure 8), a general acid to protonate the leaving group (Z-H) and the presence of one or more positively charged groups to stabilise the transition state (Galburt and Stoddard, 2002) (Figure 8). The role of a general base can be fulfilled both by metal cofactors and/or protein side chains. Divalent metal ions can also serve as Lewis acids (Y in the scheme), stabilizing the pentacoordinate intermediate.

One of the most relevant nucleophiles in the enzyme catalysis is the hydroxide ion ( $\text{OH}^-$ ) and in order to produce it from water, enzymes generally use one or more metal cofactors. The  $\text{pK}_a$  of a water molecule coordinated to one or more divalent metal ions is reduced, thereby generating a hydroxide ion in close proximity to the protein active site (Gerlt, 1993 and 1992). Moreover, the DNA coordination to one or more divalent metal ions reduces or even eliminates the electrostatic repulsion between the water derived nucleophile and the negatively charged DNA backbone (Galburt and Stoddard, 2002), (Figure 8). Upon the  $\text{OH}^-$  attack, an OH-DNA pentacoordinate intermediate is formed (Guthrie, 1997). Creation and stabilisation of this intermediate within the nucleases' active site permits the enzymatic reaction to greatly exceed in speed compared to non-enzymatic reactions, where stabilisation of such intermediates is not possible (Jencks, 1981).

Endonucleases have been described to require one, two or three bound metals per active site. Exceptions are some ribonucleases, which cleave the phosphodiester bonds in RNA. They

do not use metal ions as cofactors, but instead utilize only protein side chains to provide the necessary positive charges, and to act as proton donors and acceptors (Fersht, 1999). Possible explanation for the lack of metals in some ribonucleolytic active sites could be the relative instability of the RNA phosphodiester bond, as compared to that of the DNA. And in that regard RNA cleaving enzymes have an easier job than the DNA nucleases (Galburt and Stoddard, 2002).

### 3.1 Endonucleases and one, two or three metal ion catalysis

#### *Three metal ions mechanism*

Three metal ions coordination has been described for the BER enzyme Endonuclease IV (Hosfield *et al.*, 1999) and suggested as being part of the UVDE repair mechanism (Paspaleva *et al.*, 2007). Endonuclease IV uses a cluster of  $Zn^{2+}$  ions, while UVDE most likely utilizes three  $Mn^{2+}$  ions. Classically, the three metal ion mechanism proceeds as follows: the metal ion in site I facilitates the formation of the nucleophilic hydroxide, whereas the metal ions in site II and III contribute mainly to stabilisation of the transition state and assist the leaving group. A detailed scheme of the Endo IV repair mechanism and the involvement of the  $Zn^{2+}$  cluster in the DNA cleavage can be found in chapter 4.2 of this introduction.

#### *Two metal ions mechanism*

Two metal ions catalysis has been described for the restriction enzymes *Bam*HI ( $Ca^{2+}$ ) (Viadiu and Aggarwal, 1998), *Bgl*I ( $Ca^{2+}$ ) (Newman *et al.*, 1998) and *Pvu*II ( $Ca^{2+}$ ) (Horton and Cheng, 2000). One of the metal ions was shown to be responsible for lowering the  $pK_a$  of a neighbouring water molecule, facilitating its deprotonation. Both metal ions are required to stabilise the transition state: one interacts with the oxygen of the  $OH^-$  nucleophile, and the other with the leaving group (Pingoud *et al.*, 2005). As seen in the crystal structures of *Bam*HI, *Bgl*I and *Pvu*II, protein side chains assist the activation of the catalytic water and in that regard the presence of a third divalent metal ion is no longer needed.

#### *One metal ion mechanism*

Enzymes can even use only one metal ion for their DNA cleaving function. One-metal ion catalysis has been described for the restriction nucleases: *Eco*RI ( $Mn^{2+}$ ) (McClarín *et al.*, 1986), *Eco*O109I ( $Mn^{2+}$ ) and *Hinc*II ( $Ca^{2+}$ ) (Etzkorn and Horton, 2004). For all these nucleases the DNA hydrolysis reaction starts with the binding of a water molecule to the metal ion near the active site, resulting in its suitable positioning for the subsequent nucleophilic attack of the phosphorus atom. The water molecule is polarised through its coordination to the divalent metal ion coordination to the divalent metal ion, but mainly to protein side chains.

In summary, a number of nucleases utilize metal ions for their catalytic mechanism. The most commonly used divalent metal ions are  $Mg^{2+}$ ,  $Mn^{2+}$ ,  $Ca^{2+}$  and to lesser extend  $Zn^{2+}$ ,  $Cu^{2+}$  and

Co<sup>2+</sup> (Pingoud *et al.*, 2005). Although the different nucleases might vary in respect to how many divalent metal ions are required for the hydrolysis of the DNA phosphodiester backbone, the underlying mechanism is the same. The divalent ions activate the nucleophilic water molecule and/or serve as efficient cofactors to stabilise the pentacoordinate transition state.

## 4. BASE EXCISION REPAIR

Base Excision Repair (BER) is a cellular mechanism, which targets a large variety of alterations of the DNA bases (Wilson and Thompson, 1997). The first step of BER involves removal of the modified base from the deoxyribose by DNA glycosylase. DNA glycosylases bind specifically to a target base and hydrolyse the N-glycosylic bond, releasing the damaged base, while keeping the DNA backbone intact. This step is referred to as damage specific, since each DNA glycosylase recognises one type of base alteration. Upon the glycosylase activity, an apurinic/apyrimidinic site is formed, which is substrate for AP-endonucleases. In the damage-general step of BER AP endonuclease cleaves the phosphodiester backbone 5' to the abasic site, resulting in the formation of a 3'-hydroxyl and 5' abasic sugar phosphate. The 5' abasic sugar phosphate is then removed with the help of exonuclease, or by specific DNA-deoxyribosephosphodiesterase (dRpase). The sequential action of DNA glycosylases, AP endonucleases and exonucleases result in the creation of single nucleotide gap, which is filled by DNA polymerase and finally the resulting nick is sealed by DNA ligase.

A number of DNA glycosylases also possess DNA lyase activity, which in contrast to the AP endonucleases involves cleavage of the DNA backbone 3' to the AP-site. The activity of the combined glycosylases/lyases also results in creation of a single nucleotide gap, which can be filled by a DNA polymerase and sealed by a ligase.

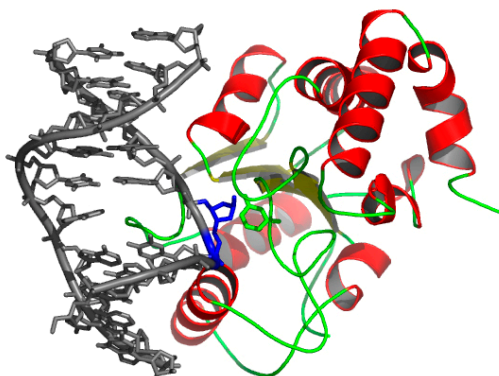
### 4.1 Glycosylases

#### *Uracil-DNA glycosylase*

A Uracil in DNA is a result of cytosine deamination or can be product of dUTP incorporation. It is specifically removed by Uracil-DNA Glycosylase (UDG), which catalyses the hydrolysis of the N-glycosylic bond between the uracil (U) and the DNA sugar, leaving an AP site. UDGs are widely distributed small monomeric proteins (Friedberg *et al.*, 2005). They classically do not require metal cofactors for their activity, although the DNA binding of some UDGs is strongly stimulated by Mg<sup>2+</sup> ions (Kavli *et al.*, 2002).

*E. coli* UDG is the founding representative of the class 1 UDG enzymes (Lindahl, 1974). The members of this family are able to remove uracil bases from both single and double strand DNA. Crystal structures are available for several members of the class 1 UDG family, including the *E. coli* homologue. The structure of *E. coli* UDG (Pearl, 2000) revealed an  $\alpha/\beta$  fold, with a central four stranded  $\beta$  sheet. The co-crystal with an uracil containing DNA (Parikh *et al.*,

2000) (Figure 9) elucidated the *E. coli* UDG damage recognition mechanism. The enzyme was seen to kink the DNA at the lesion site and flip the abnormal nucleotide into an uracil specific pocket, where catalysis takes place. The shape of the pocket provides selection against purines. Furthermore, a tyrosine residue (Tyr66) blocks the accommodation of 5-methylated pyrimidines (thymines). A conserved asparagine (Asn123) assists in discrimination between cytosine and uracil by forming hydrogen bonds only with uracil.



**Figure 9.** Structure of the *E. coli* Uracil-DNA glycosylase after hydrolysis of the N-glycosylic bond.

The uracil base is colored in green and the abasic site is shown in blue.

DNA glycosylases catalyse the nucleophilic displacement of a damaged base from the DNA using a water derived hydroxyl group. The reaction catalysed by the *E. coli* UDG involves the utilisation of Asp64 as a general base, which activates the water molecule for the subsequent attack of the N-glycosyl bond (Drohata *et al.*, 1999). In addition, the closely located His187 assists in the electrostatic stabilisation of the uracil in the intermediate reaction phase. Since both the nucleophile creation and the intermediate stabilisation are performed by the UDG protein side chains, there is no actual need for divalent metal cofactors, which in other enzymes have been described to perform the same role.

#### *T4 DNA glycosylase (T4 endo V)*

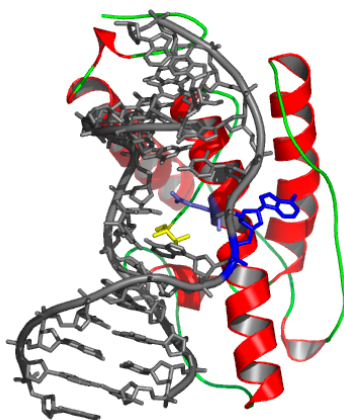
The bacteriophage T4 encodes an enzyme for the specific removal of CPDs from the DNA. T4 endo V is 18 kDa *cis-syn* cyclobutane pyrimidine dimer specific glycosylase, with an additional AP lyase activity (Minton *et al.*, 1975). The enzyme cleaves the N-glycosyl bond of the 5'-pyrimidine of the dimer and has no requirements for divalent cations or other cofactors (Friedberg *et al.*, 1971). Upon the N-glycosyl bond cleavage, T4 endo V incises the DNA backbone 3' to the abasic site (lyase activity).

The T4 endonuclease V has been also shown to cleave the *trans-syn* CPD in a double-stranded DNA at a significant rate, although at least 100 times slower than the *cis-syn* dimer (Smith and Taylor, 1993).

Homologues of T4 endo V have been found in a limited number of eubacteria: in some *Prochlorococcus marinus* strains, *Brucella*, *Bordetella* as well as in *Sinorhizobium medicae* and *Pasteurella multocida* (Goosen and Moolenaar, 2007).

The X-ray crystal structures of the T4 endo V wild type (Morikawa *et al.*, 1992) and a mutant (Nakabeppu *et al.*, 1982) protein have been determined. The wild type T4 endo V is seen as a single domain structure of three  $\alpha$ -helices, with the amino terminal region situated between helix one (H1) and helix three (H3), and in close proximity to the proposed catalytic residue (Glu23). The region accommodating Glu23 and the N-terminal Thr2 (Met1 is removed from the enzyme *in vivo*), which acts as the active site nucleophile is positioned in the centre of a groove, with dimensions just large enough to accommodate a single strand form of the DNA (Morikawa *et al.*, 1992). The structures of three catalytically inactive mutants (E23Q, E23D, and R3Q) show an almost identical peptide backbone structure to the wild type.

A 2.75 Å co-crystal structure of an inactive T4-endonuclease V mutant (E23Q) with pyrimidine dimer-containing DNA duplex has also been determined (Golan *et al.*, 2006). The structure of the enzyme is remarkably unchanged in the co-crystal and the DNA is bent at the dimer site by 60° (Figure 10). Upon DNA binding, the active site residues: Glu23 and the N-terminal Thr2  $\alpha$ -amino group protrude into the DNA stack.



**Figure 10.** Structure of T4 endo V with DNA.

The N-terminal  $\alpha$ -amino group, which acts as the active site nucleophile is colored in yellow. Glu23 is shown in violet and the flipped out adenine, opposite the scissile base, in blue.

There is no direct contact of the enzyme with the thymines of the pyrimidine dimer, but there are extensive interactions between protein side chains and the deformed DNA backbone,

near the dimer site. The Glu23 carboxyl side chain is positioned near the sugar, and the Thr2  $\alpha$ -amino group is close to the 5' thymine of the CPD, correctly positioned to act as the nucleophile in the glycosylase reaction. The adenine opposite the target thymine in the photodimer is flipped out of the DNA structure and bound into a pocket on the side of the enzyme. The adenine in the pocket does not form hydrogen bonds with any protein residues, suggesting that the major forces holding the adenine in place are relatively non-specific van der Waals interactions. However, some protein residues do assist in the stabilisation of the flipped conformation, since the extrahelical adenine is seen to be stacked between Tyr21 and Arg22.

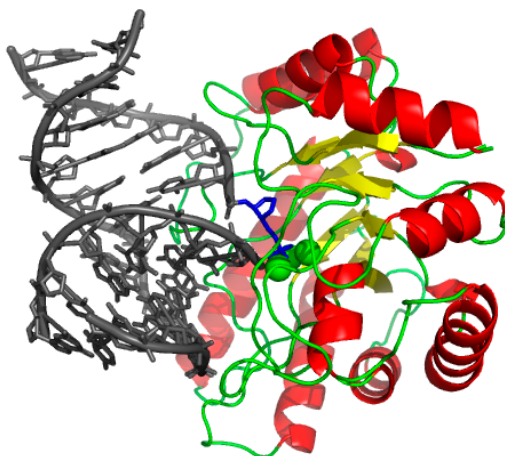
Upon cleavage of the N-glycosylic bond of the 5'T of the CPD dimer, the T4 endo V catalyses a  $\beta$ -elimination reaction that cleaves DNA on the 3' side of the CPD (Vassilyev *et al.*, 1995). Mutational studies showed that the main residue responsible for the enzyme AP lyase activity is Glu23, since mutation in this amino acid completely abolishes the DNA lyase activity (Hori *et al.*, 1993).

## 4.2 AP endonucleases

AP endonucleases are metalloenzymes introducing a nick directly 5' to an AP site, as part of the second step of BER. They are ubiquitous and here only selected examples of well characterised AP endonucleases will be described.

### *Endonuclease IV (Nfo) family of AP endonucleases*

Nfo (endonuclease IV) is an endonuclease, which nicks AP sites in *E. coli*. The AP sites are cleaved directly 5' to the lesion, leaving a hydroxyl group at the 3' terminus and a deoxyribose 5'-phosphate at the 5' terminus (Ljungquist, 1977). Endonuclease IV was shown not only to

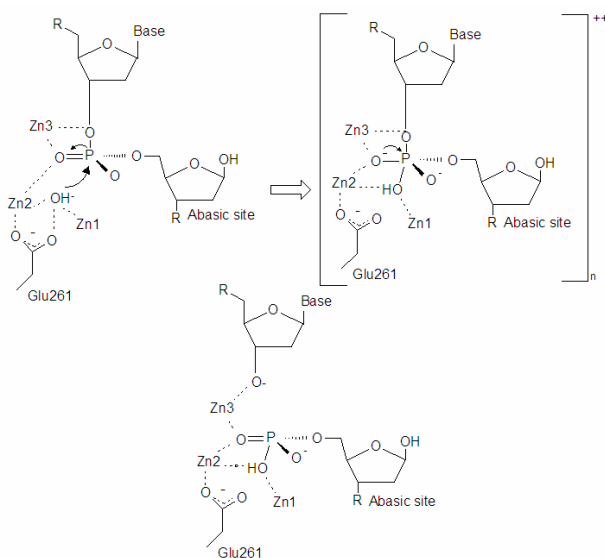


**Figure 11.** Crystal structure of Endo IV with an abasic site containing DNA.

The flipped abasic site is shown in blue and the three Zn<sup>2+</sup> ions are colored in green.

process abasic site lesions but also to act on a variety of oxidative damages in DNA (Souza *et al.*, 2006). The crystal structure of the *E. coli* Endo IV (Hosfield *et al.*, 1999) revealed a TIM barrel fold and the presence of three  $Zn^{2+}$  ions buried in a deep, crescent-shaped protein groove. The presence of the metal cofactors and the close position of the C- and the N- termini classify Endo IV in the TIM barrel fold family of “divalent metal dependent enzymes”. The high resolution (1.5 Å) crystal structure of Endo IV with its AP-DNA complex (Figure 11) shows severe bending of the DNA backbone of approximately  $90^\circ$ , which promotes double nucleotide flipping and positioning of the extrahelical AP site in an active site pocket, which houses the cluster of three  $Zn^{2+}$  ions. This pocket has been proposed to be the main reason for the Endo IV selectivity for abasic sites, since it sterically excludes normal nucleotides (Hosfield *et al.*, 1999).

In the suggested mechanisms for the abasic site cleavage, the DNA backbone is incised with the help of the trinuclear zinc cluster, with all three  $Zn^{2+}$  ions participating in the catalysis (Figure 12). The role of Zn1 was seen in the crystal structure to be predominantly in activating a water molecule and stabilising the important nucleophile ( $OH^-$ ), needed for the nucleophilic attack on the target P-O DNA bond. Zn2 together with Glu261 also assists in orienting and activating the attacking nucleophile. The role of Zn3 is mainly in neutralizing the charge of the phosphate group, rendering the phosphorus atom susceptible to nucleophilic substitution. Upon nucleophile attack by the bridging hydroxide, the reaction proceeds through an intermediate pentacoordinate transition state (Figure 12). All  $Zn^{2+}$  ions participate in the stabilisation of the intermediate step. When the transition state collapses to the reaction products, the negative charge at  $O3'$  is neutralised by the interaction with Zn3, while Zn2 and Zn1 assist in the stabilisation of the leaving (abasic site containing) group.



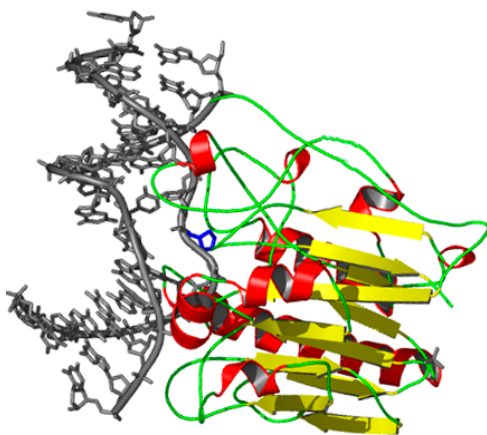
**Figure 12.** Structure based catalytic mechanism of Endo IV.

Another member of the endonuclease IV family is the *S. cerevisiae* homologue Apn1. Apn1 shows 41 % sequence identity to the *E. coli* Endo IV and also uses three  $Zn^{2+}$  for its function.

### APE1

In humans, the majority of the 5' AP endonuclease activity is provided by APE1 (Dempfle *et al.*, 1991), also known as HAP1, REF1 and APEX. Unlike Endo IV, APE1 is an  $Mg^{2+}$  dependent nuclease.

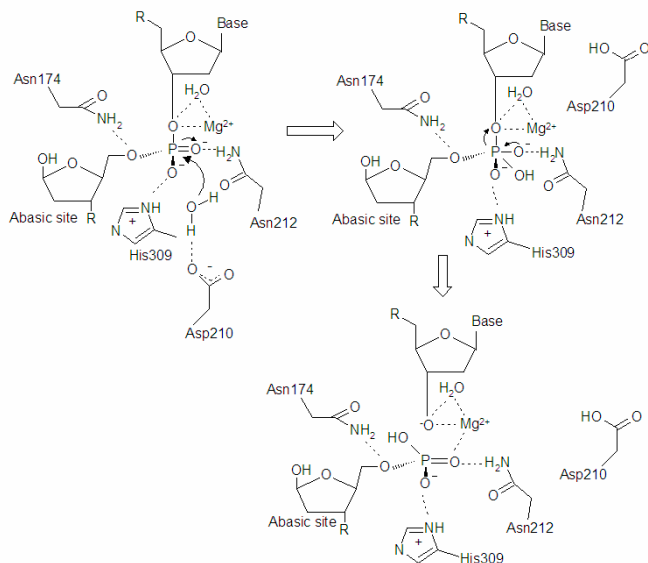
Crystal structures of the human APE1 bound to a synthetic abasic site-containing DNA, both with and without the divalent  $Mg^{2+}$  ion, show how APE1 recognises abasic sites and cleaves the target bond (Beernink *et al.*, 2001). In the co-crystal structures, APE1 was seen to flip the abasic nucleotide in an active site protein pocket. The enzyme-DNA interface was seen to include both DNA strands, although the interaction with the AP-DNA strand is prenominal (Figure 13).



**Figure 13.** Structure of APE1. The flipped AP site is shown in blue.

Within the pocket the AP-DNA substrate is oriented with the help of one divalent metal ion ( $Mg^{2+}$ ) and APE1 active-site residues, while Asp210 is aligned for activating the nucleophilic hydroxyl (Figure 14). Upon the nucleophile attack, an intermediate complex is formed and stabilised by Asn212, Asn174, His309 and the single  $Mg^{2+}$ . Subsequently, the collapse of the transition state leads to the cleavage of the scissile P-O3' bond, with the O3' leaving group stabilised by the metal ion (Figure 14).

To summarise, in the catalytic reaction of APE1 only one divalent metal cofactor ( $Mg^{2+}$ ) is used, and unlike endonuclease IV, the magnesium ion is not directly involved in the activation of the catalytic water. This role is performed by Asp210, while the role of the metal cofactor is mainly in the orientation of the target P-O bond, stabilisation of the pentacoordinate intermediate and the O3' leaving group.



**Figure 14.** Structure based reaction mechanism for the phosphodiester bond cleavage of APE1.

## 5. NUCLEOTIDE EXCISION REPAIR

Nucleotide excision repair (NER) is an important mechanism for the removal of a large variety of structurally unrelated DNA lesions such as intra-strand crosslinks (like cisplatin adducts), bulky mono-adducts (like N-2-acetylaminofluorene) and UV induced photoproducts (Van Houten *et al.*, 2005). The basic mechanism of NER is conserved in pro- and eukaryotes, but the proteins involved are different. NER can be divided into several steps: damage recognition, dual incision on both sides of the lesion, removal of the damaged oligonucleotide and finally gap filling and sealing of the resulting nick.

Since the most crystal structures are available for the bacterial NER proteins, here only this NER system will be discussed.

### 5.1 Bacterial NER

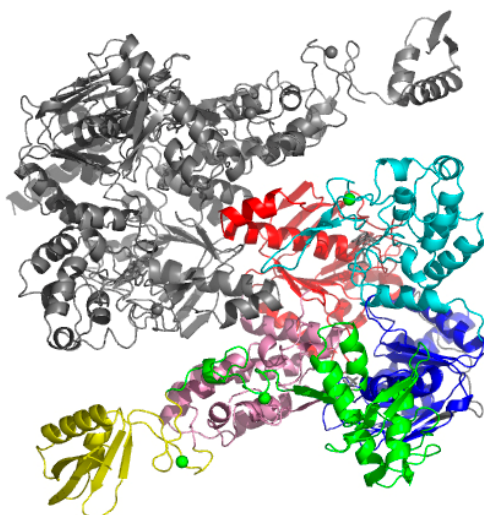
The process of NER is performed in *E. coli* by the UvrABC system, which consists of four Uvr proteins: UvrA, UvrB, UvrC, and DNA helicase II (also known as UvrD) (Truglio *et al.*, 2006). The UvrA and UvrB enzymes associate in solution, forming an UvrA<sub>2</sub>-UvrB<sub>2</sub> complex (Malta *et al.*, 2007). Upon encountering DNA damage, the UvrA dimer is released, while the UvrB dimer remains bound to the DNA (Orren and Sancar, 1990). The process of damage recognition, the UvrA dissociation and the formation of the pre-incision complex are dependent on ATP binding and hydrolysis (Goosen and Moolenaar, 2001). Subsequently, UvrC binds to the UvrB<sub>2</sub>-DNA preincision complex, displacing one of the UvrB units. UvrC introduces two nicks: first 4

nucleotides from the 3' side of the DNA damage and then 8 nucleotides from the 5' side (Truglio *et al.*, 2005), (Verhoeven *et al.*, 2000). Subsequently, the DNA fragment containing the damage is removed by the help of UvrD, which also removes the UvrC protein. The resulting gap is then filled in using DNA polymerase I and sealed by DNA ligase.

### *UvrA*

UvrA plays an important role in the *E. coli* NER system initiating the damage recognition and is capable of binding to damaged DNA, even in the absence of the other NER proteins (Truglio *et al.*, 2006). At physiological concentrations UvrA is a dimer and its dimerisation, as well as its interactions with DNA, is promoted by ATP binding (Mazur and Grossman, 1991). *E. coli* UvrA has two ATPase domains, both belonging to the ABC (ATP-binding cassette) family of ATPases (Myles and Sancar, 1991), (Gorbalenya, 1990). Each of the ABC ATPase domains consists of a Walker A, a Walker B and an ABC signature motif (Myles and Sancar, 1991).

Recently, a 3.2 Å UvrA crystal structure from *Bacillus stearothermophilus* has been reported (Pakotiprapha *et al.*, 2008) (Figure 15). It revealed an UvrA dimer, with each monomer containing six domains: ATP binding domains I and II, signature domains I and II (with the sequence LSGGQ), the UvrB-binding domain and an insertion domain. The ATP-binding domain I and the signature domain I are located in the N-terminal half, while the C-terminal part contains the ATP-binding domain II and the signature domain II. ATP binds to both the ATP binding domain in the N-terminus and to the signature motif in the C-terminus. The main difference



**Figure 15.** 3.2 Å structure of UvrA from *Bacillus stearothermophilus*.

The UvrB binding domain is in yellow. The signature motif I is shown in pink and the signature motif II in light blue. The ATP binding site I is colored red and the ATP binding site II is represented in dark blue. The insertion domain is colored in green. The Zn atoms are shown as green balls.

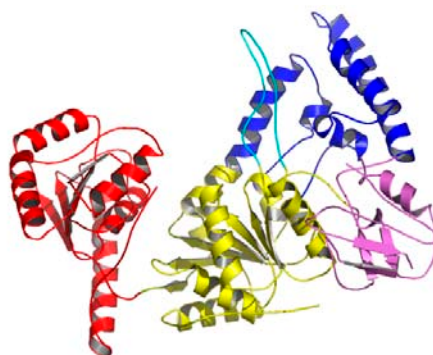
between the N- and the C-terminal part of UvrA is the presence of the UvrB binding domain (residues 118 – 256) and the insertion domain (residues 287 – 398) in the N-terminal half, which both contains Zn modules. Deletion mutations in the UvrB binding domain revealed that residues 113 – 245 are crucial for the UvrB-UvrA interactions and based on the UvrB structure it was proposed that they make a contact with domain 2 of UvrB (Pakotiprapha *et al.*, 2009).

Anomalous diffraction data obtained from the UvrA structure showed the presence of three Zn atoms. Zn1 is located between the signature domain I and the UvrB binding domain. Zn2 is situated between the signature domain I and the so called insertion domain. Zn3 connects the signature domain II to the dimer interface. The Zn modules do not show a classical Zn finger structure and are proposed to play a structural role (Pakotiprapha *et al.*, 2007).

Although the published *B. stearothermophilus* UvrA structure does not contain DNA, based on the sequence conservation and the positive charge Pakotiprapha *et al.* suggested the potential DNA interacting region to be located in the concave side of the UvrA dimer. Unfortunately, the process of UvrA damage recognition is not yet clarified, although it has been suggested that Lys680 and Arg691 might be important for DNA binding, since replacement of these side chains with alanine decrease the UvrA-DNA binding 3 - 37 folds (Croteau *et al.*, 2008).

### *UvrB*

UvrB is a central component of the bacterial NER system participating in the important step of damage recognition (Hsu *et al.*, 1995). *E. coli* UvrB consists of 673 amino acids and has a molecular weight of 76 kDa. So far three crystal structures of the UvrB protein are available – two from the thermophile organism *Thermus thermophilus* (Machius *et al.*, 1999), (Nakagawa *et al.*, 1999) and one from a different thermophile *Bacillus caldotenax* (Theis *et al.*, 1999) (Figure 16). The three UvrB structures share a similar overall architecture. UvrB consists of five domains: 1a,



**Figure 16.** Crystal structure of the DNA Repair enzyme UvrB from *B. caldotenax*.

The hairpin is represented in a light blue color, domain 1a is in yellow, domain 1b is in blue, domain 2 is in pink and domain 3 in red.

1b, 2, 3, 4 (Figure 16). Domain 4 is missing in the crystal structure of *B. caldotenax*, because it is disordered. UvrB domain 4 has been shown to be important for the UvrC binding (Moolenaar *et al.*, 1995) and dimerization (Malta *et al.*, 2007). An UvrB C-terminal truncation (UvrB\*), lacking domain 4, was shown to form a less stable dimer (Verhoeven *et al.*, 2002). X-ray (Sohi *et al.*, 2000) and NMR structures (Alexandrovich *et al.*, 1999) of this UvrB domain are available, and both models agree with it forming a dimer in solution.

An interesting feature of the UvrB structure is the presence of a  $\beta$ -hairpin, situated between domains 1a and 1b (Skorvaga *et al.*, 2002). Mutational analysis of the *E. coli* hairpin has shown that aromatic residues (Tyr92 and Tyr93), situated at its bottom, are involved in preventing binding to undamaged DNA (Moolenaar *et al.*, 2001).

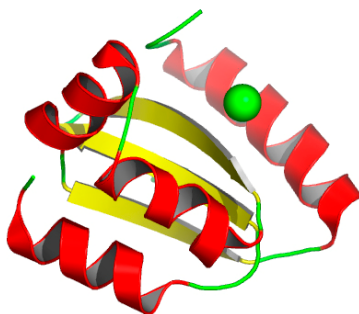
The importance of the  $\beta$ -hairpin for the DNA binding was confirmed in the recently published structures of UvrB in a complex with DNA having a 3' overhang (Truglio *et al.*, 2006). The crystal structure of *B. caldotenax* UvrB bound to this DNA substrate revealed its structure to be highly similar to the native enzyme, the main difference being in the  $\beta$ -hairpin conformation. In the co-crystal of *B. caldotenax* UvrB one of the DNA strands was seen to pass behind the  $\beta$ -hairpin (Truglio *et al.*, 2006) and one base from this strand to be flipped out and accommodated into a small hydrophobic pocket. This pocket is too small to accommodate big DNA distortions (for example a cholesterol lesion) and only a planar molecule can fit in. In that aspect, the flipped nucleotide is unlikely to represent the damaged one. A drawback of this crystallized complex is that UvrB it is not bound to a specific DNA damage, but is in a complex with double strand-single strand DNA junction. In that regard, it is difficult to state if the DNA strand observed behind the hairpin represents the damaged- or the non-damaged strand.

Nucleotide flipping by UvrB was demonstrated by using the fluorescent base analogue 2-aminopurine (Malta *et al.*, 2006). It was shown that the base at the 3' side of the lesion is inserted into a protein pocket. Furthermore, by using menthol modification covalently attached to a 2-amino purine residue, Malta *et al.* (2008) outlined that the damaged base itself does not change position upon UvrB binding and keeps its intra-helical conformation. Although the exact location of the menthol was not clarified, based on its hydrophobic nature the authors predicted that it also remains buried inside the DNA helix. A model was proposed in which the damaged strand passes behind the  $\beta$ -hairpin and the presence of a damage prevents the translocation of this strand behind the hairpin, which limits the base flipping to the nucleotide 3' to the damage.

### *UvrC*

UvrC is a 68 kDa nuclease, which mediates both the 3' and the 5' incision reactions on damaged DNA upon binding to the UvrB-DNA pre-incision complex (Sancar and Rupp, 1983). Site-directed mutagenesis showed that the 3' incision active site is located in the N-terminal half of the protein (Verhoeven *et al.*, 2000), while the site, which carries out the 5' cut is situated in the C-terminal half (Lin and Sancar, 1992). The 3' incision occurs prior to the 5' one (Truglio *et al.*, 2005).

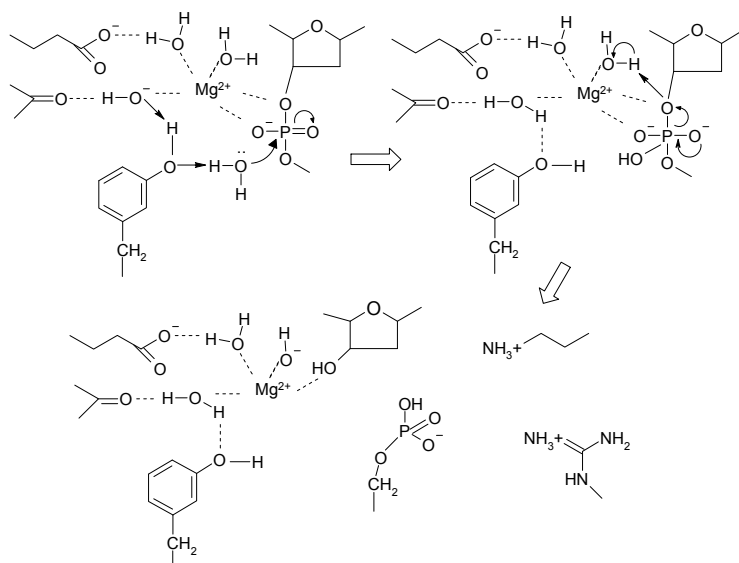
Multiple sequence analysis of the N-terminal UvrC part revealed that the catalytic domain is a member of the GIY-YIG family of homing endonucleases (Kowalski *et al.*, 1999). The crystal structures of the N-terminal 3' catalytic domains from *B. caldotenax* and *Thermotoga maritima* (Figure 17) have been solved (Truglio *et al.*, 2005).



**Figure 17.** Structure of the C-terminal part of UvrC from *T. maritima*.

The bound metal is shown as green sphere.

Despite soaking and co-crystallization attempts, no divalent metal was detected in the *B. caldotenax* UvrC structure (Truglio *et al.*, 2005). The *T. maritima* UvrC did reveal one divalent metal, coordinated by Glu76 and five water molecules, however due to its low occupancy the exact nature of the cofactor was not elucidated. Soaking experiments performed in the presence of  $\text{MnCl}_2$  suggested that  $\text{Mn}^{2+}$  can bind within that site, although the fact that  $\text{Mg}^{2+}$  also stimulates

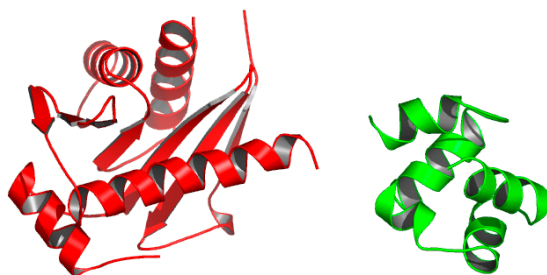


**Figure 18.** Proposed mechanism for UvrC catalysis.

UvrC activity suggests that the cofactor can also be  $Mg^{2+}$ . Although Glu76 was seen as the only protein residue bound directly to the metal, the water molecules coordinating the metal formed additional contacts with the protein side chains. Mutation of Glu76 proved the crucial role of this residue for the UvrC function, since the alanine substitution resulted in a catalytically inactive UvrC enzyme (Truglio *et al.*, 2005).

Interestingly, from the structural data the molecule acting as the general acid in the catalytic process was not found and it was proposed that one of the metal coordinated water molecules performs this function (Truglio *et al.*, 2005). Furthermore, Tyr29, which is situated in close proximity to the divalent cation (4.3 Å), was suggested to act as the general base by accepting a proton from the nucleophilic water molecule and simultaneously transferring its proton to the metal-bound hydroxide (Figure 18).

The C-terminal 219 amino acids of the UvrC from *T. maritima* were crystallised (Karakas *et al.*, 2007) and revealed an endonuclease domain responsible for the 5' incision event (residues 341 – 494), followed by a pair of helix–hairpin–helix (HhH) motifs (residues 497 – 557), that form an (HhH)<sub>2</sub> domain implicated in DNA binding (Figure 19). A very short linker (residues 495 and 496) connects the two domains.



**Figure 19.** Structure of the C-terminal part of UvrC from *T. maritima*.

The endonuclease and the (HhH)<sub>2</sub> domain are shown in red and green, respectively.

The core of the UvrC 5' endonuclease domain has a fold belonging to the Rnase H family of enzymes. Divalent cations ( $Mg^{2+}$  or  $Mn^{2+}$ ) are required by Rnase H-like enzymes to bind to their substrate and catalyze the nucleotidyl transfer reactions (Krakowiak *et al.*, 2002).  $MnCl_2$  soaking experiments with UvrC revealed one catalytic  $Mn^{2+}$ , coordinated by His488. Two catalytic aspartates (Asp367 and Asp429) interact indirectly with the cation through water molecules. Since the H488A mutant is still able to perform the 5' incision, albeit with a reduced rate, it is unlikely that this residue is solely responsible for metal coordination. Based on the similarity of UvrC C-terminal domain to the Rnase H-like enzymes a two metal ions catalysis was proposed for the UvrC 5' incision, where the binding of the second metal depends greatly on the presence of a nucleic acid substrate (Karakas *et al.*, 2007).

## 6. UVDE

### 6.1 Discovery and initial characterisation of UVDE

Analysis of the radiation sensitivity of the yeast *Schizosaccharomyces pombe* led to the identification of a group of mutants, which phenotype could not be explained by defects in the NER system (Phipps *et al.*, 1985). Later it was found that *S. pombe* rad mutants, which are deficient in NER, are only moderately UV sensitive and still capable of efficient removal of both CPD and 6-4PP UV lesions via a light independent pathway (McCready *et al.*, 1993), (Sidik *et al.*, 1992). Based on these data it was proposed that *S. pombe* possesses a novel excision repair pathway for removal of UV induced lesions (Yonemasu *et al.*, 1997). The first biochemical evidence that this pathway exists was the identification of an ATP-independent UV damage-dependent endonuclease in *S. pombe* cell extracts (Bowman *et al.*, 1994). The enzyme was initially named *S. pombe* DNA endonuclease (SPDE), but was renamed later to UV damage endonuclease (UVDE). The initial biochemical analysis revealed that UVDE (also designated Uvelp) cleaves the DNA phosphodiester backbone directly 5' to both CPD and 6-4PP lesions (Bowman *et al.*, 1994). The isolation of the *uve1* gene (Takao *et al.*, 1996) and the biochemical characterization of a truncated UVDE protein (Kaur *et al.*, 1998) revealed a much broader substrate specificity of the *S. pombe* UVDE than originally suspected.

### 6.2 Biochemical characterisation of *S. pombe* UVDE

The full length *uve1* gene encodes a 599 amino acid protein of approximately 68.8 kDa, which when overexpressed in either *E. coli* or *S. cerevisiae* cells, yields a relatively unstable protein that rapidly loses activity (Kaur *et al.*, 1998). In an attempt of obtaining an *S. pombe* UVDE protein with better stability several affinity tagged truncations were purified and their activity towards UV lesions analysed. Most of the truncated proteins contained N-terminal deletions with different length with a sharp decline in activity for truncations longer than 232 amino acids (Takao *et al.*, 1996). Truncations of the C-terminal part of UVDE were shown to have more drastic effect, since even a relatively short deletion of 35 residues abrogates activity (Takao *et al.*, 1996). Based on these findings, the essential region for the activity of the *S. pombe* UVDE was proposed to include the C-terminal two thirds of the protein. Kaur *et al.* conducted a detailed biochemical characterisation of a highly stable 228 residue N-terminal truncation of UVDE. This protein was shown to have broad substrate specificity and being active in a large range of salt concentrations with a pH optimum between pH 6.0 and 6.5.

### 6.3 Activity of UVDE on UV induced lesions

The initial studies of *S. pombe* UVDE demonstrated that this enzyme is capable of recognising both *cis-syn* CPDs and 6-4 PPs (Bowman *et al.*, 1994). Later, the incision activity of UVDE was determined on a series of UV induced lesions: *cis-syn* CPD (89 % incision), *trans-syn* I CPD (75 %), *trans-syn* II CPD (75 %), 6-4 PP (71 %) and the Dewar 6-4 PP isomer (83 %), (Avery

*et al.*, 1999). Each of the UV photoproducts listed above, causes different distortion when incorporated into DNA, showing the capacity of UVDE to recognise a range of structurally unrelated damages.

#### **6.4 Activity of UVDE on abasic sites**

UVDE was shown to cleave DNA substrates containing abasic sites in a manner similar to the AP endonucleases (Tanihigashi *et al.*, 2006), introducing a nick directly 5' to the abasic site and producing 3'-hydroxyl and 5'-deoxyribose phosphate (Avery *et al.*, 1999). However, unlike the AP endonucleases, UVDE introduces a second nick one nucleotide away from the lesion (Avery *et al.*, 1999). The presence of a second nick, however, is highly dependent on the DNA sequence. The UVDE activity on the abasic site as well as its actual incision position will be addressed in further details in Chapter 4 of this thesis. Abasic site substrates containing either A or G opposite to the lesions are processed by *N. crassa* UVDE with equal efficiency, whereas lesions with C or T opposite are cleaved less efficiently (Kano *et al.*, 1999), suggesting that UVDE interacts with the undamaged DNA strand. The same authors demonstrated that introduction of UVDE into *E. coli* mutants lacking both of its major AP endonucleases (exonuclease III and endonuclease IV) gave the host cells a resistance to abasic sites introducing agents such as methyl methane sulfonate (MMS) and butylhydroperoxide, however to lesser extent than the AP endonucleases. Based on these findings, a new role has been proposed for UVDE (Kanno *et al.*, 1999) as being part of an alternative repair system for abasic sites. Our findings, however, question this notion, as it will be shown in Chapter 5, that not all UVDE enzymes recognise the abasic site efficiently. Furthermore, the activity of *S. pombe* UVDE on the abasic site is strictly sequence dependent.

#### **6.5 Recognition of base-base mismatches**

*In vitro* studies (Kaur *et al.*, 1999) established that to various extents, UVDE also recognises and cleaves single base mismatches. Moreover, it was shown by the same authors that mutants lacking the *uve1* gene display spontaneous mutator phenotype, underlining their hypothesis that UVDE might play a role in mismatch repair (Kunz *et al.*, 2001). The sites of UVDE cleavage were mapped to be 5' immediately to the impaired base and/or one nucleotide further away, dependent on the specific mismatch. The cleavage efficiency was also shown to vary significantly (from 5 % to 40 %), depending on the nature of the base mispair (Kaur *et al.*, 1999). The low incision efficiency of UVDE questions the role of this enzyme as an alternative pathway for mismatch repair. The observed mutator phenotype of the *uve1* mutants might also be caused by reduced removal of abasic sites.

UVDE has also been shown to incise duplex DNA containing small loops and hairpins, albeit again with a very low efficiency. A study performed by Kaur and Doetsch (Kaur and Doetsch, 2000) demonstrated that UVDE incises DNA containing small loops of 2-4 nucleotides in length. The position of the nick, however, is not as expected directly 5' to the loop but several nucleotides away. Similar behaviour has been observed also on a Pt-crosslink, where the incision

position is also not directly adjacent to the damage. What determines the position of the UVDE cut will be further elucidated in Chapter 4. Larger loops of 6-8 nucleotides in length are not a substrate for UVDE. DNA with palindromic insertions that could form base-paired hairpin structures is also not nicked by UVDE.

### 6.7 Distribution of UVDE homologues

The distribution of UVDE has been recently reviewed by Goosen and Moolenaar (2007). Homologues of the UVDE protein have been identified in 4 archaeobacteria and in 30 eubacterial species, as well as in several fungi. A number of closely related *Bacillus* species (*Bacillus anthracis*, *Bacillus cereus* and *Bacillus thuringiensis*) contain even two UVDE homologues. It has been suggested (Goosen and Moolenaar, 2007) that one of the UVDE gene products is expressed in the vegetative state, while the other is specifically expressed upon spore germination. Another possibility is that the two UVDE enzymes exhibit difference in their substrate specificities, since in spores a specific spore photoproduct is formed upon UV irradiation. The only species outside the *Bacillales* that also contain two copies of UVDE is *Clostridium beyerinckii*, which is also a spore-forming bacterium (Goosen and Moolenaar, 2007). A homologous protein to the *S. pombe* UVDE has been also found in the *Deinococcus radiodurans*. Unlike the  $\Delta$ uvrA strains of *E. coli*, the  $\Delta$ uvrA strains of *D. radiodurans* exhibit nearly wild-type levels of resistance to UV light (Earl *et al.*, 2002).

Interestingly, the N-terminal one third of the UVDE protein is not conserved in the eukaryotic homologues and completely absent in the bacterial ones. The N-terminal region is highly hydrophilic, as it includes many charged amino acids. The C-terminal part of the protein varies in length and also contains a number of charged residues. The function of these extensions is as yet unknown.

### 6.8 The UVDE repair pathway

A simple model has been proposed for the steps following the nicking activity of UVDE (Yasui and McCready, 1998). The repair process was suggested to be completed by simply removing a segment of DNA containing the UV photoproduct with a 5' - 3' exonuclease (or with an endonuclease cutting 3' to the UV photoproduct site), filling the gap by a polymerase and sealing the resulting nick by a DNA ligase (Doetsch 1995). Epistatic analyses have implicated several gene products to be involved in the UVDE-initiated DNA repair, including the *S. pombe* FEN-1 (flap endonuclease) homolog and the exonuclease Rad2p (Yoon *et al.*, 1999). The first complete *in vitro* reconstitution of the UVDE pathway has been given by Alleva *et al.* (2000). Two distinct pathways have been proposed. In the first pathway upon the 5' nick introduced by UVDE, the damage is removed by the 5' to 3' exonuclease activity of Rad2p, resulting in formation of a gap. The gap is subsequently filled by polymerase  $\delta$  in a presence of PCNA (Proliferating Cell Nuclear Antigen) and RFC (Replication Factor C) and the resulting nick sealed by a ligase. Alternatively, the DNA polymerase can first extend the free 3' end produced after the UVDE

nicking activity, thus creating a misplaced 5' flap. This flap structure is a substrate for the FEN-1 flap-endonuclease activity. The resulting nick is again sealed by a ligase.

The broad substrate specificity of UVDE poses the question why bacterial species have both UVDE and the NER repair pathways. Although the substrate specificities of both systems overlap to some extent, NER distinguish itself with broader substrate specificity than UVDE. On the CPD lesion, however, the UVDE is beneficial since it repairs this lesion with high affinity.

**REFERENCES**

- Alexandrovich A, Sanderson MR, Moolenaar GF, Goosen N and Lane AN, NMR assignments and secondary structure of the UvrC binding domain of UvrB, *FEBS Lett.* **451** (1999), pp. 181-185
- Alleva JL, Zuo S, Hurwitz J and Doetsch PW, In vitro reconstitution of the *Schizosaccharomyces pombe* alternative excision repair pathway, *Biochemistry* **39** (2000), pp. 2659-2666
- Avery AM, Kaur B, Taylor JS, Mello JA, Essigmann JM and Doetsch PW, Substrate specificity of ultraviolet DNA endonuclease (UVDE/Uve1p) from *Schizosaccharomyces pombe*, *Nucleic Acids Res.* **27** (1999), pp. 2256-2264
- Beernink T, Segelke W, Hadi Z, Erzberger P, Wilson M 3<sup>rd</sup> and Rupp B, Two divalent metal ions in the active site of a new crystal form of human apurinic/apyrimidinic endonuclease, Ape1: implications for the catalytic mechanism, *J. Mol. Biol.* **307** (2001), pp. 1023-1034
- Bowman KK, Sidik K, Smith CA, Taylor JS, Doetsch PW and Freyer GA, A new ATP-independent DNA endonuclease from *Schizosaccharomyces pombe* that recognizes cyclobutane pyrimidine dimers and 6-4 photoproducts, *Nucleic Acids Res.* **22** (1994), pp. 3026-3032
- Carell T, Burgdorf LT, Kundu LM and Cichon M, The mechanism of action of DNA photolyases, *Curr. Opin. Chem. Biol.* **5** (2001), pp. 491-498
- Carey F and Sundberg R, *Advanced organic chemistry*, 3<sup>rd</sup> edition. Plenum Publishing Corporation, New York (2000)
- Croteau DL, DellaVecchia MJ, Perera L and Van Houten B, Cooperative damage recognition by UvrA and UvrB: identification of UvrA residues that mediate DNA binding, *DNA Repair (Amst.)* **7** (2008), pp. 392-404
- Daniels DS, Woo TT, Luu KX, Noll DM, Clarke ND, Pegg AE and Tainer JA, DNA binding and nucleotide flipping by the human DNA repair protein AGT, *Nat. Struct. Mol. Biol.* **1** (2004), pp. 714-720
- David S, Goodsell, *The Molecular Perspective: Ultraviolet Light and Pyrimidine Dimers*, *Oncologist* **6** (2001), pp. 298-299
- Demple B, Herman T and Chen S, Cloning and expression of APE, the cDNA encoding the major human apurinic endonuclease: definition of a family of DNA repair enzymes, *Proc. Natl. Acad. Sci. USA.* **88** (1991), pp. 11450-11454
- Demple B and Linn S, 5,6-saturated thymine lesions in DNA: production by ultraviolet light or hydrogen peroxide, *Nucleic Acids Res.* **10** (1982), pp. 3781-3789
- Demple B, Sedgwick B, Robins P, Totty N, Waterfield MD and Lindahl T, Active site and complete sequence of the suicidal methyltransferase that counters alkylation mutagenesis, *Proc. Natl. Acad. Sci. USA.* **82** (1985), pp. 2688-2692
- Doetsch PW, What's old is new: an alternative DNA excision repair pathway, *Trends Biochem. Sci.* **20** (1995), pp. 384-386
- Drohat AC, Jagadeesh J, Ferguson E and Stivers JT, Role of electrophilic and general base catalysis in the mechanism of *Escherichia coli* uracil DNA glycosylase, *Biochemistry* **8** (1999), pp. 11866-11875
- Earl AM, Rankin SK, Kim KP, Lamendola ON and Battista JR, Genetic evidence that the uvsE gene product of *Deinococcus radiodurans* R1 is a UV damage endonuclease, *J. Bacteriol.* **184** (2002), pp. 1003-1009
- Essen LO and Klar T, Light-driven DNA repair by photolyases, *Cell. Mol. Life Sci.* **63** (2006), pp. 1266-1277
- Etzkorn C and Horton C, Ca<sup>2+</sup> binding in the active site of HincII: implications for the catalytic mechanism, *Biochemistry* **43** (2004), pp. 13256-13270

- Fersht A, Structure and Mechanism in Protein Science: a guide to enzyme catalysis and protein folding, W. H. Freeman and Co., New York (1999)
- Friedberg C and King J, Dark repair of ultraviolet-irradiated deoxyribonucleic acid by bacteriophage T4: purification and characterization of a dimer-specific phage-induced endonuclease, *J. Bacteriol.* **106** (1971), pp. 500-507
- Friedberg C, Walker C, Siede W, Wood D, Schultz A and Ellenberger T, DNA Repair and Mutagenesis, 2<sup>nd</sup> Edition, ASM Press (2005)
- Galburt E and Stoddard B, Catalytic mechanisms of restriction and homing endonucleases, *Biochemistry* **41** (2002), pp. 13851-13860
- Gerlt J, Mechanistic principles of enzyme-catalysed cleavage of phosphodiester bonds, Nucleases, 2<sup>nd</sup> edition, Cold Spring Harbor Laboratory Press (1993)
- Gerlt J, Phosphate ester hydrolysis, *Enzymes* **20** (1992), pp. 95-139
- Golan G, Zharkov O, Grollman P, Dodson L, McCullough K, Lloyd S and Shoham G, Structure of T4 pyrimidine dimer glycosylase in a reduced imine covalent complex with abasic site-containing DNA, *J. Mol. Biol.* **362** (2006), pp. 241-258
- Goodtzova K, Kanugula S, Edara S, Pauly GT, Moschel RC and Pegg AE, Repair of O6-benzylguanine by the *Escherichia coli* Ada and Ogt and the human O6-alkylguanine-DNA alkyltransferases, *J. Biol. Chem.* **272** (1997), pp. 8332-8338
- Goosen N and Moolenaar GF, Repair of UV damage in bacteria, *DNA Repair (Amst.)* **7** (2007), pp. 353-379
- Goosen N and Moolenaar GF, Role of ATP hydrolysis by UvrA and UvrB during NER, *Res. Microbiol.* **152** (2001), pp. 401-409
- Gorbalenya AE and Koonin EV, Superfamily of UvrA-related NTP-binding proteins. Implications for rational classification of recombination/repair systems, *J. Mol. Biol.* **213** (1990), pp. 583-591
- Guo H, Yang H, Mockler TC and Lin C, Regulation of flowering time by *Arabidopsis* photoreceptors, *Science* **279** (1998), pp. 1360-1363
- Guthrie J, Hydration and dehydration of phosphoric acid derivatives: Free energies of formation of the pentacoordinate intermediates for phosphate ester hydrolysis and of monomeric monophosphate, *J. Chem. Soc.* **99** (1997), pp. 3391-4001
- Hitomi K, Nakamura H, Kim S.-T, Mizukoshi T, Ishikawa T, Iwai S and Todo T, Role of two histidines in the (6-4) photolyase reaction, *J. Biol. Chem.* **276** (2001), pp. 10103-10109
- Hori N, Doi T, Karaki Y, Kikuchi M, Ikenhara M and Ohtsuka E, Participation of glutamic acid 23 of T4 endonuclease V in the beta-elimination reaction of an abasic site in synthetic duplex DNA, *Nucleic Acid Res.* **20** (1993), pp. 4761-4764
- Horton JR and Cheng X, PvuII endonuclease contains two calcium ions in active sites, *J. Mol. Biol.* **300** (2000), pp. 1049-1056
- Hosfield J, Guan Y, Haas J, Cunningham P and Tainer A, Structure of the DNA repair enzyme endonuclease IV and its DNA complex: double-nucleotide flipping at abasic sites and three-metal-ion catalysis, *Cell* **98** (1999), pp. 397-408
- Hsu DS, Kim ST, Sun Q and Sancar A, Structure and function of the UvrB protein, *J. Biol. Chem.* **270** (1995), pp. 8319-8327
- Husain I, Griffith J and Sancar A, Thymine dimers bend DNA, *Proc. Natl. Acad. Sci. USA.* **85** (1988), pp. 2558-2562

- Jackson PE, Hall CN, O'Connor PJ, Cooper DP, Margison GP and Povey AC, Low O<sup>6</sup>-alkylguanine DNA-alkyltransferase activity in normal colorectal tissue is associated with colorectal tumours containing a GC→AT transition in the K-ras oncogene, *Carcinogenesis* **18** (1997), pp. 1299-1302
- Jencks W, How does a reaction choose its mechanism? *Chem. Soc. Rev.* **10** (1981), pp. 345-375
- Jing Y, Kao JF and Taylor JS, Thermodynamic and base-pairing studies of matched and mismatched DNA dodecamer duplexes containing *cis-syn*, (6-4) and Dewar photoproducts of TT, *Nucleic Acids Res.* **26** (1998), pp. 3845-3853
- Kanno S, Iwai S, Takao M and Yasui A, Repair of apurinic/apyrimidinic sites by UV damage endonuclease; a repair protein for UV and oxidative damage, *Nucleic Acids Res.* **27** (1999), pp. 3096-3103
- Kanugula S, Goodtzova K and Pegg AE, Probing of conformational changes in human O<sup>6</sup>-alkylguanine-DNA alkyl transferase protein in its alkylated and DNA-bound states by limited proteolysis, *Biochem J.* **32** (1998), pp. 545-545
- Karakas E, Truglio JJ, Croteau D, Rhau B, Wang L, Van Houten B and Kisker C, Structure of the C-terminal half of UvrC reveals an RNase H endonuclease domain with an Argonaute-like catalytic triad, *EMBO J.* **26** (2007), pp. 613-622
- Kaur B, Avery AM and Doetsch PW, Expression, purification, and characterization of ultraviolet DNA endonuclease from *Schizosaccharomyces pombe*, *Biochemistry* **37** (1998), pp. 11599-11604
- Kaur B and Doetsch PW, Ultraviolet damage endonuclease (Uve1p): a structure and strand-specific DNA endonuclease, *Biochemistry* **39** (2000), pp. 5788-5796
- Kaur B, Fraser JL, Freyer GA, Davey S and Doetsch PW, A Uve1p-mediated mismatch repair pathway in *Schizosaccharomyces pombe*, *Mol. Cell. Biol.* **19** (1999), pp. 4703-4710
- Kavli B, Sundheim O, Akbari M, Otterlei M, Nilsen H, Skorpen F, Aas PA, Hagen L, Krokan HE and 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.* **277** (2002), pp. 39926-39936
- Kim JK and Choi BS, The solution structure of DNA duplex-decamer containing the (6-4) photoproduct of thymidyl (3'→5') thymidine by NMR and relaxation matrix refinement, *Eur. J. Biochem.* **228** (1995), pp. 849-854
- Kim ST, Heelis PF, Okamura T, Hirata Y, Mataga N and Sancar A, Determination of rates and yields of interchromophore (folate→flavin) energy transfer and intermolecular (flavin→DNA) electron transfer in *Escherichia coli* photolyase by time-resolved fluorescence and absorption spectroscopy, *Biochemistry* **30** (2001), pp. 11262-11270
- Kim ST, Heelis PF and Sancar A, Energy transfer (deazaflavin→FADH<sub>2</sub>) and electron transfer (FADH<sub>2</sub>→TT) kinetics in *Anacystis nidulans* photolyase, *Biochemistry* **31** (1992), pp. 11244-11248
- Kobayashi K, Kanno S, Smit B, van der Horst GT, Takao M and Yasui A, Characterization of photolyase/blue-light receptor homologs in mouse and human cells, *Nucleic Acids Res.* **26** (1998), pp. 5086-5092
- Kobayashi Y, Ishikawa T, Hirayama J, Daiyasu H, Kanai S, Toh H, Fukuda I, Tsujimura T, Terada N, Kamei Y, Yuba S, Iwai S and Todo T, Molecular analysis of zebrafish photolyase/cryptochrome family: two types of cryptochromes present in zebrafish, *Genes Cells* **5** (2000), pp. 725-738
- Kowalski JC, Belfort M, Stapleton MA, Holpert M, Dansereau JT, Pietrokovski S, Baxter SM and Derbyshire V, Configuration of the catalytic GIY-YIG domain of intron endonuclease I-TevI: coincidence of computational and molecular findings, *Nucleic Acids Res.* **27** (1999), pp. 2115-2125
- Krakowiak A, Owczarek A, Koziol-Kiewicz M and Stec WJ, Stereochemical course of *Escherichia coli* RNase H, *Chembiochem.* **3** (2002), pp. 1242-1250
- Kunz C and Fleck O, Role of the DNA repair nucleases Rad13, Rad2 and Uve1 of *Schizosaccharomyces pombe* in mismatch correction, *J. Mol. Biol.* **313** (2001), pp. 241-253

- Lee JH, Hwang GS, Kim JK and Choi BS, The solution structure of DNA decamer duplex containing the Dewar product of thymidyl (3→5') thymidine by NMR and full relaxation matrix refinement, *FEBS Lett.* **428** (1998), pp. 269-274
- Lin JJ and Sancar A, Active site of (A)BC excinuclease I. Evidence for 5' incision by UvrC through a catalytic site involving Asp399, Asp438, Asp466, and His538 residues, *J. Biol. Chem.* **267** (1992), pp. 17688-17692
- Lindhahl T and Nyberg B, Heat-induced deamination of cytosine residues in deoxyribonucleic acid, *Biochemistry* **13** (1974), pp. 3405-3410
- Lindhahl T, An N-glycosidase from *Escherichia coli* that releases free uracil from DNA containing deaminated cytosine residues, *Proc. Natl. Acad. Sci. USA.* **71** (1974), pp. 3649-3653
- Liu H, Hewitt SR and Hays JB, Antagonism of ultraviolet-light mutagenesis by the methyl-directed mismatch-repair system of *Escherichia coli*, *Genetics* **154** (2000), pp. 503-512
- Ljungquist S, A new endonuclease from *Escherichia coli* acting at apurinic sites in DNA, *J. Biol. Chem.* **252** (1977), pp. 2808-2814
- Machius M, Henry L, Palnitkar M and Deisenhofer J, Crystal structure of the DNA nucleotide excision repair enzyme UvrB from *Thermus thermophilus*, *Proc. Natl. Acad. Sci. USA.* **96** (1999), pp. 11717-11722
- Malta E, Moolenaar GF and Goosen N, Base flipping in nucleotide excision repair, *J. Biol. Chem.* **281** (2006), pp. 2184-2194
- Malta E, Moolenaar GF and Goosen N, Dynamics of the UvrABC nucleotide excision repair proteins analyzed by fluorescence resonance energy transfer, *Biochemistry* **46** (2007), pp. 9080-9088
- Malta E, Verhagen CP, Moolenaar GF, Filippov DV, van der Marel GA and Goosen N, Functions of base flipping in *E. coli* nucleotide excision repair, *DNA Repair (Amst.)* **7** (2008), pp. 1647-1658
- Mazur SJ and Grossman L, Dimerisation of *E. coli* UvrA and its binding to undamaged and ultraviolet light damaged DNA, *Biochemistry* **30** (1991), pp. 4431-4443
- McClarin A, Frederick C, Wang C, Greene P, Boyer W, Grable J and Rosenberg M, Structure of the DNA-*EcoRI* endonucleases recognition complex at 3 Å resolution, *Science* **234** (1986), pp. 1526-1534
- McCready S, Carr AM and Lehmann AR, Repair of cyclobutane pyrimidine dimers and 6-4 photoproducts in the fission yeast *Schizosaccharomyces pombe*, *Mol. Microbiol.* **10** (1993), pp. 885-890
- Mees A, Klar T, Gnau P, Hennecke U, Eker AP, Carell T and Essen LO, Crystal structure of a photolyase bound to a CPD-like DNA lesion after in situ repair, *Science* **306** (2004), pp. 1789-1793
- Minton K, Durphy M, Taylor R and Friedberg C. The ultraviolet endonuclease of bacteriophage T4: further characterization. *J. Biol. Chem.* **250** (1975), pp. 2823-2829
- Mitchell DL, Jen J and Cleaver JE, Sequence specificity of cyclobutane pyrimidine dimers in DNA treated with solar (ultraviolet B) radiation, *Nucleic Acids Res.* **20** (1992), pp. 225-229
- Mitchell DL and Nairn RS, The biology of the (6-4) photoproduct, *Photochem. Photobiol.* **49** (1989), pp. 805-819
- Moolenaar GF, Franken KL, Dijkstra DM, Thomas-Oates JE, Visse R, van de Putte P and Goosen N, The C-terminal region of the UvrB protein of *Escherichia coli* contains an important determinant for UvrC binding to the preincision complex but not the catalytic site for 3'-incision, *J. Biol. Chem.* **270** (1995), pp. 30508-30515
- Moolenaar GF, Höglund L and Goosen N, Clue to damage recognition by UvrB: residues in the beta-hairpin structure prevent binding to non-damaged DNA, *EMBO J.* **20** (2001), pp. 6140-6149
- Moore MH, Gulbis JM, Dodson EJ, Demple B and Moody PCE, Crystal structure of a suicidal DNA repair protein: the Ada O<sup>6</sup>-methylguanine-DNA methyltransferase from *E. coli*, *EMBO J.* **13** (1994), pp. 1495-1501

- Morikawa K, Matsumoto O, Tsujimoto M, Katayanagi K, Ariyoshi M, Doi T, Ikehara M, Inaoka T and Ohtsuka E, X-ray structure of T4 endonuclease V: an excision repair enzyme specific for a pyrimidine dimer, *Science* **256** (1992), pp. 523-526
- Myles GM and Sancar A, Isolation and characterization of functional domains of UvrA, *Biochemistry* **30** (1991), pp. 3834-3840
- Nakabeppu Y, Yamashita K and Sekiguchi M, Purification and characterization of normal and mutant forms of T4 endonuclease V, *J. Biol. Chem.* **257** (1982), pp. 2556-2562
- Nakagawa N, Sugahara M, Masui R, Kato F, Fukuyama K and Kuramitsu S, Crystal structure of *Thermus thermophilus* HB8 UvrB protein, a key enzyme of nucleotide excision repair, *J. Biochem.* **126** (1999), pp. 986-990
- Nakajima S, Sugiyama M, Iwai S, Hitomi K, Otoshi E, Kim S.-T, Jiang C.-Z, Todo T, Britt AB and Yamamoto K, Cloning and characterization of a gene (*UVR3*) required for photorepair of 6-4 photoproducts in *Arabidopsis thaliana*, *Nucleic Acids Res.* **26** (1998), pp. 638-644
- Newman M, Lunnen K, Wilson G, Greci J, Schildkraut I and Phillips SE, Crystal structure of restriction endonuclease *BglI* bound to its interrupted DNA recognition sequence, *EMBO J.* **17** (1998), pp. 5466-5476
- Orren DK and Sancar A, Formation and enzymatic properties of the UvrB-DNA complex, *J. Biol. Chem.* **265** (1990), pp. 15796-15803
- Paalman SR, Sung C and Clarke ND, Specificity of DNA repair methyltransferases determined by competitive inactivation with oligonucleotide substrates: evidence that *Escherichia coli* Ada repairs O<sup>6</sup>-methylguanine and O<sup>6</sup>-methylthymine with similar efficiency, *Biochemistry* **37** (1997), pp. 11118-11124
- Pakotiprapha D, Inuzuka Y, Bowman BR, Moolenaar GF, Goosen N, Jeruzalmi D and Verdine GL, Crystal structure of *Bacillus stearothermophilus* UvrA provides insight into ATP-Modulated dimerization, UvrB interaction, and DNA binding, **29** *Mol. Cell* (2008), pp. 122-133
- Parikh SS, Walcher G, Jones GD, Slupphaug G, Krokan HE, Blackburn GM and Tainer JA, Uracil-DNA glycosylase-DNA substrate and product structures: conformational strain promotes catalytic efficiency by coupled stereoelectronic effects, *Proc. Natl. Acad. Sci. USA.* **97** (2000), pp. 5083-5088
- Park H, Zhang K, Ren Y, Nadji S, Sinha N, Taylor JS and Kang C, Crystal structure of a DNA decamer containing a *cis-syn* thymine dimer, *Proc. Natl. Acad. Sci. USA.* **99** (2000), pp. 5965-5970
- Paspaleva K, Thomassen E, Pannu NS, Iwai S, Moolenaar GF, Goosen N and Abrahams JP, Crystal structure of the DNA repair enzyme ultraviolet damage endonuclease, *Structure* **15** (2007), pp. 1316-1324
- Pearl L, Structure and function in the uracil-DNA glycosylase superfamily, *Mutat. Res.* **460** (2000), pp. 165-181
- Pearlman DA, Holbrook SR, Pirkle DH and Kim SH, Molecular models for DNA damaged by photoreaction, *Science* **227** (1985), pp. 1304-1308
- Phipps J, Nasim A and Miller DR, Recovery, repair, and mutagenesis in *Schizosaccharomyces pombe*, *Adv. Genet.* **23** (1985), pp. 1-72
- Pingoud A, Fuxreiter M, Pingoud V and Wende W, Type II restriction endonucleases: structure and mechanism, *Cellular and Molecular Life Sciences (CMLS)* **62** (2005), pp. 685-707
- Samson L, The suicidal DNA repair methyltransferases of microbes, *Mol. Microbiol.* **6** (1992), pp. 825-831
- Sancar A and Rupp WD, A novel repair enzyme: UVRABC excision nuclease of *Escherichia coli* cuts a DNA strand on both sides of the damaged region, *Cell* **33** (1983), pp. 249-260
- Sancar GB, Enzymatic photoreactivation: 50 years and counting, *Mutat. Res.* **451** (2000), pp. 25-37
- Sedgwick B, Bates PA, Paik J, Jacobs SC and Lindahl T, Repair of alkylated DNA, *DNA Repair (Amst.)* **6** (2007), pp. 429-442

- Sidik K, Lieberman HB and Freyer GA, Repair of DNA damaged by UV light and ionizing radiation by cell-free extracts prepared from *Schizosaccharomyces pombe*, *Proc. Natl. Acad. Sci. USA.* **89** (1992), pp. 1212-1216
- Skorvaga M, Theis K, Mandavilli BS, Kisker C and Van Houten B, The  $\beta$ -hairpin motif of UvrB is essential for DNA binding, damage processing and UvrC-mediated incisions, *J. Biol. Chem.* **277** (2002), pp. 1553-1559
- Smith A and Taylor S, Preparation and characterization of a set of deoxyoligonucleotide 49-mers containing site-specific *cis-syn*, *trans-syn*-I, (6-4) and Dewar photoproducts of thymidylyl (3'-5') thymidine, *Biol. Chem.* **268** (1993), pp. 11143-11151
- Sohi M, Alexandrovich A, Moolenaar G, Visse R, Goosen N, Vernede X, Fontecilla-Camps JC, Champness J and Sanderson MR, Crystal structure of *Escherichia coli* UvrB C-terminal domain, and a model for UvrB-uvrC interaction, *FEBS Lett.* **465** (2000), pp. 161-164
- Souza L, Eduardo R, Padula M and Leitao C, Endonuclease IV and exonuclease III are involved in the repair and mutagenesis of DNA lesions induced by UVB in *Escherichia coli*, *Mutagenesis* **21** (2006), pp. 125-130
- Takao M, Yonemasu R, Yamamoto K and Yasui A, Characterization of a UV endonuclease gene from the fission yeast *Schizosaccharomyces pombe* and its bacterial homolog, *Nucleic Acids Res.* **24** (1996), pp. 1267-1271
- Tanihigashi H, Yamada A, Igawa E and Ikeda S, The role of *Schizosaccharomyces pombe* DNA repair enzymes Apn1p and Uve1p in the base excision repair of apurinic/aprimidinic sites, *Biochem. Biophys. Res. Commun.* **347** (2006), pp. 889-894
- Taylor JS, Garrett DS and Cohrs MP, Solution-state structure of the Dewar pyrimidinone photoproduct of thymidylyl-(3'-5')-thymidine, *Biochemistry* **27** (1988), pp. 7206-7215
- Theis K, Chen PJ, Skorvaga M, Van Houten B and Kisker C, Crystal structure of UvrB, a DNA helicase adapted for nucleotide excision repair, *EMBO J.* **18** (1999), pp. 6899-6907
- Todo T, Kim S.-T, Hitomi K, Otoshi E, Inui T, Morioka H, Kobayashi H, Ohtsuka E, Toh H and Ikenaga M, Flavin adenine dinucleotide as a chromophore of the *Xenopus* (6-4) photolyase, *Nucleic Acids Res.* **257** (1997), pp. 764-768
- Todo T, Ryo H, Yamamoto K, Toh H, Inui T, Ayaki H, Nomura T and Ikenaga M, Similarity among the *Drosophila* (6-4) photolyase, a human photolyase homolog, and the DNA photolyase-blue-light photoreceptor family, *Science* **272** (1996), pp. 109-112
- Todo T, Functional diversity of the DNA photolyase/blue light receptor family, *Mutat. Res.* **434** (1999), pp. 89-97
- Truglio JJ, Croteau DL, Van Houten B and Kisker C, Prokaryotic nucleotide excision repair, the UvrABC system, *Chem. Rev.* **106** (2006), pp. 233-252.
- Truglio JJ, Rhau B, Croteau DL, Wang L, Skorvaga M, Karakas E, DellaVecchia MJ, Wang H, Van Houten B and Kisker C, Structural Insights into the first incision reaction during nucleotide excision repair, *EMBO J.* **24** (2005), pp. 885-894
- Van Houten B, Croteau DL, DellaVecchia MJ, Wang H and Kisker C, 'Close-fitting sleeves': DNA damage recognition by the UvrABC nuclease system, *Mutat. Res.* **577** (2005), pp. 92-117
- Vassilyev G, Kashiwagi T, Mikami Y, Ariyoshi M and Iwai S, 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
- Verhoeven EE, van Kesteren M, Moolenaar GF, Visse R and Goosen N, Catalytic sites for 3' and 5' incision of *Escherichia coli* nucleotide excision repair are both located in UvrC, *J. Biol. Chem.* **275** (2000), pp. 5120-5123

- Verhoeven EE, Wyman C, Moolenaar GF and Goosen N, The presence of two UvrB subunits in the UvrAB complex ensures damage detection in both DNA strands, *EMBO J.* **21** (2002), pp. 4196-4205
- Viadiu H and Aggarwal AK, The role of metals in catalysis by the restriction endonuclease *Bam*HI, *Nat. Struct. Biol.* **5** (1998), pp. 910-916
- Wang CI and Taylor JS, Site-specific effect of thymine dimer formation on dAn.dTn tract bending and its biological implications, *Proc. Natl. Acad. Sci. USA.* **85** (1991), pp. 2558-2562
- Weber S, Light-driven enzymatic catalysis of DNA repair: a review of recent biophysical studies on photolyase, *Biochim. Biophys. Acta* **1707** (2005), pp. 1-23
- Westheimer FH, Why nature chose phosphates, *Science* **235** (1987), pp. 1173-1178
- Wilson M and Thompson H, Life without DNA repair, *Proc. Natl. Acad. Sci. USA.* **94** (1997), pp. 12754-12757
- Worthington E, Kavakli I, Berrocal-Tito G, Bondo B and Sancar A, Purification and characterization of three members of the photolyase/cryptochrome family blue-light photoreceptors from *Vibrio cholerae*, *J. Biol. Chem.* **278** (2003), pp. 39143-39154
- Yasui A, Eker AP, Yasuhira S, Yajima H, Kobayashi T, Takao M and Oikawa A, A new class of DNA photolyases present in various organisms including aplacental mammals, *EMBO J.* **13** (2001), pp. 6143-6151
- Yasui A and McCready SJ, Alternative repair pathways for UV-induced DNA damage, *Bioessays* **20** (1998), pp. 291-297
- Yonemasu R, McCready SJ, Murray JM, Osman F, Takao M, Yamamoto K, Lehmann AR and Yasui A, Characterization of the alternative excision repair pathway of UV-damaged DNA in *Schizosaccharomyces pombe*, *Nucleic Acids Res.* **25** (1997), pp. 1553-1558
- Yoon JH, Lee CS, O'Connor TR, Yasui A and Pfeifer GP, The DNA damage spectrum produced by simulated light, *J. Mol. Biol.* **299** (2000), pp. 681-693
- Yoon JH, Swiderski PM, Kaplan BE, Takao M, Yasui A, Shen B and Pfeifer GP, Processing of UV damage in vitro by FEN-1 proteins as part of an alternative DNA excision repair pathway, *Biochemistry* **38** (1999), pp. 4809-4817
- Zhao X, Liu J, Hsu DS, Zhao S, Taylor JS and Sancar A, Reaction mechanism of (6-4) photolyase, *J. Biol. Chem.* **272** (1997), pp. 32580-32590



# Crystal structure of the DNA repair enzyme ultraviolet damage endonuclease

Chapter

2

Keti Paspaleva<sup>1</sup>, Ellen Thomassen<sup>2</sup>, Navraj S. Pannu<sup>2</sup>, Shigenori Iwai<sup>3</sup>, Geri F. Moolenaar<sup>1</sup>, Nora Goosen<sup>1</sup> and Jan Pieter Abrahams<sup>2</sup>

---

<sup>1</sup>Laboratory of Molecular Genetics, Leiden Institute of Chemistry, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands

<sup>2</sup>Department of Biophysical Structural Chemistry, Leiden Institute of Chemistry, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands

<sup>3</sup>Division of Chemistry, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan

## SUMMARY

The ultraviolet damage endonuclease (UVDE) performs the initial step in an alternative excision repair pathway of UV-induced DNA damage, nicking immediately adjacent to the 5' phosphate of the damaged nucleotides. Unique for a single-protein DNA repair endonuclease, it can detect different types of damage. Here we show that *Thermus thermophilus* UVDE shares some essential structural features with Endo IV, an enzyme from the base excision repair pathway that exclusively nicks at abasic sites. A comparison between the structures indicates how DNA is bound by UVDE, how UVDE may recognize damage, and which of its residues are involved in catalysis. Furthermore, the comparison suggests an elegant explanation of UVDE's potential to recognize different types of damage. Incision assays including point mutants of UVDE confirmed the relevance of these conclusions.

## INTRODUCTION

Guarding genetic integrity by repairing damaged DNA is one of the most fundamental processes of life. To meet this challenge, life has evolved a range of DNA repair pathways, including nucleotide excision repair (NER), base excision repair (BER), and the alternative ultraviolet damage endonuclease (UVDE) repair. NER recognizes many types of damage for which it requires a large, multiprotein complex (reviewed in Truglio *et al.*, 2006). BER employs a collection of glycosylases, each recognizing and removing a different set of modified DNA bases (reviewed in Friedberg *et al.*, 2005). The UVDE pathway distinguishes itself by recognizing and subsequently nicking DNA containing different types of damage with the single, multifunctional UVDE enzyme (Takao *et al.*, 1996). Although functionally characterized, there are no structural data of the UVDE enzyme.

There are some similarities between the BER and UVDE pathways. In the first step of BER, a glycosylase removes the damaged base, leaving just the deoxyribose. In the next step, the DNA 5' adjacent to this abasic site is nicked by an apurinic/apyrimidinic endonuclease. Subsequently, other enzymes remove the nicked residue and resynthesize the damaged strand (Friedberg *et al.*, 2005). In UVDE repair the first step is skipped, and the UVDE enzyme immediately recognizes and nicks the damaged DNA strand. Even though the UVDE enzyme is much more versatile in damage recognition than the BER endonuclease, it was proposed to have a similar topology (Aravind *et al.*, 1999). The structure of the Endo IV endonuclease of the BER pathway and its complex with abasic, nicked DNA are known. (Hosfield *et al.*, 1999).

Although the UVDE DNA repair pathway was originally thought to be specific for UV damage, enzymatic studies revealed that UVDE from *Schizosaccharomyces pombe* possesses a broader substrate specificity including pyrimidine dimers (CPD), 6-4 photoproducts (6-4PP), apurinic/apyrimidinic (AP) sites, uracil (U), dihydrouracil (DHU), and other non-UV-induced DNA adducts (Avery *et al.*, 1999; Kanno *et al.*, 1999). Biochemical and genetic analysis also suggest that UVDE may be involved in orchestrating mismatch repair in vivo (Kaur *et al.*, 1999).

The broad substrate specificity of *S. pombe* UVDE was recently reviewed by Paul W. Doetsch (Doetsch *et al.*, 2006), showing that this enzyme is also active on insertion-deletion loops, supporting the role of UVDE in mismatch repair. In conclusion, the broad substrate specificity of UVDE indicates that it recognizes a common distortion in the DNA helix rather than the chemical structure of the DNA damage.

The UVDE pathway was described for the first time in the fission yeast *S. pombe*. Homologues of UVDE are present in many fungal species but also in a number of bacteria, such as *Bacillus subtilis* and the thermophilic bacterium *Thermus thermophilus*. Overexpressed, purified full-length UVDE protein from *S. pombe* (68 kDa) is unstable (Kaur *et al.*, 1998), and a number of biochemical studies have been done using a truncated protein lacking its N-terminal 228 amino acids. The bacterial homologues of UVDE lack this N-terminal region and also the

highly charged C-terminal domain of their eukaryotic counterparts. We determined the structure of UVDE from *T. thermophilus* to answer some of the main questions in the field: what is the catalytic mechanism of UVDE and how does it recognize so many different types of damage?

## EXPERIMENTAL PROCEDURES

### Cloning, expression and purification of the UVDE homolog from *T. thermophilus*

The full-length *T. thermophilus* UVDE-coding region was initially amplified from the chromosomal DNA by PRC with primers 5'-GCTTCTCATATGATCCGCTGGGCTACCCC-3' and 5'-TCGTCTCTGCAGTCAAGGGGTTGCTAGGCCCTGCTC-3' for the 5' and the 3' ends of the fragment, respectively. Primers were designed using the genome sequence of *T. thermophilus* (Henne *et al.*, 2004). The UVDE-coding fragment was cloned into the *Nde*I and *Pst*I restriction sites of pETUVDEΔ228, a plasmid previously made in our lab to overproduce a truncated *S. pombe* UVDE protein (unpublished data). Cloning into this pET plasmid allowed for a T7 promoter-driven expression of the UVDE protein with a fusion of ten histidines. The ten His residues are attached to the N-terminal part of the protein by a short, nine amino acid linker containing a factor Xa cleavable site. The *T. thermophilus* UVDE expression vector (pUD24) was then transformed into *E. coli* BL21 (Studier *et al.*, 1990). The UVDE protein was purified from cells of a 2 l culture, harvested 2 h after induction by IPTG, and lysed by sonication in 6 ml lysis buffer (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 10 mM β-mercaptoethanol, 10 % glycerol, 1 % Triton X-100). The lysate was separated into soluble and insoluble fractions by centrifugation at 37,000 rpm for 30 min. The supernatant was loaded on a HiTrap-chelating column, which was equilibrated with buffer A (20 mM Tris-HCl (pH 7.5), 500 mM NaCl, 10 % glycerol, 10 mM β-mercaptoethanol) containing 20 mM imidazole and the protein was eluted with a 20 – 250 mM gradient of imidazole in buffer A. Pooled fractions of the *T. thermophilus* UVDE were loaded on a Resource Q column equilibrated with 20 mM Tris (pH 7.5), 10 % glycerol. The protein was eluted with a 0 – 1 M NaCl gradient in 20 mM Tris-HCl (pH 7.5). Finally, the UVDE-containing fractions were loaded on a Nap5 gel-filtration column (Amersham), and equilibrated in 20 mM Tris-HCl (pH 7.5), 150 mM NaCl. The protein fractions, which showed high purity were used in *in vitro* assays and crystallization trials.

For crystallization purposes, a selenomethionine (SeMet)-substituted UVDE protein was expressed in a methionine auxotrophic derivative of BL21 (Sohi *et al.*, 2000). The SeMet-labelled UVDE was purified by the same procedure as the non-labelled protein and showed identical enzymatic activity as the non-labelled protein. Mutants E176A and Y105A containing alanine substitutions were constructed by site-directed mutagenesis using PCR and purified like the wild-type UVDE protein. All point mutations in the *T. thermophilus* UVDE resulted in proteins showing the same elution/purification profiles as the wild-type enzymes on ion-exchange and size-exclusion columns, suggesting they were properly folded.

### **Cloning, expression, and purification of the UVDE homolog from *S. pombe***

In this study, we used an N-terminal truncation of the full-length *S. pombe* UVDE protein. The first 223 amino acids were removed as described before (Kaur *et al.*, 1998). All UVDE mutants were constructed by PCR and verified by sequencing for the absence of additional PCR-induced mutations. The wild-type  $\Delta 228$ -UVDE and mutant proteins were further purified on a HiTrap-chelating column, and hydroxyapatite and P11 phosphate cellulose columns. Detailed description of  $\Delta 228$ -UVDE protein purification will be given elsewhere (unpublished data). All *S. pombe* UVDE mutants showed the same elution/purification profiles as the wild-type enzymes.

### **DNA substrates**

The DNA substrates used in this study are 30 bp substrates containing either a CPD or a 6-4PP adduct in the sequence 5'-CTCGTCAGCATCTTCATCATACAGTCAGT-3', the *TT* representing the position of the UV lesion. The oligonucleotides containing CPD or 6-4PP lesions were synthesized as described (Iwai, 2006).

### **Incision assay**

The DNA substrates were labelled at the 5' side of the top strand using polynucleotide kinase as described (Verhoeven *et al.*, 2002). The DNA substrates (0.2 nM) were incubated with 5 nM UVDE in 20  $\mu$ l reaction mix (20 mM HEPES (pH 6.5), 100 mM NaCl, 1 mM  $MnCl_2$ ). After 15 min incubation at 30°C for *S. pombe* UVDE or 55°C for the *T. thermophilus* protein, the reactions were terminated by adding 3  $\mu$ l EDTA/SDS (0.33 M EDTA, 3.3% SDS) and 2.4  $\mu$ l glycogen (4  $\mu$ g/ $\mu$ l) followed by ethanol precipitation. The incision products were visualized on a 15 % denaturing polyacrylamide gel.

### **Crystallization**

The purified protein in 20 mM Tris (pH 7.5), 150 mM NaCl, 10 % glycerol was dialyzed against 1 $\times$  PBS (137 mM NaCl, 10 mM phosphate, 2.7 mM KCl (pH 7.4)) and concentrated to 3-5 mg/ml by centrifugation using an Ultrafree filter device (Millipore). The protein was stored at 4°C till further use. After 1 week, plate-like crystals appeared in the Eppendorf tube and grew to a size of 0.1  $\times$  0.1  $\times$  0.01 mm. Crystals were transferred to 25 % (v/v) glycerol in crystallization buffer prior to data collection.

### **Crystallographic data collection and processing**

A MAD data set was collected on beamline BM14 at the European Synchrotron Radiation Facility (ESRF) at a wavelength of 0.97800 Å (peak), 0.97850 Å (inflection point), and 0.91840 Å (high-energy remote) using the anomalous signal of the selenium atoms. The crystal was flash-frozen and kept at 100K during data collection and 360 images were collected with a rotation angle of 1°. Reflections were integrated with MOSFLM (Leslie, 1999) and merged with SCALA (Evans, 1993) from the CCP4 suite (CCP4, 1994). For data statistics, see Table 1.

### Structure solution and refinement

A strong nonorigin peak in the Patterson map indicated the presence of noncrystallographic translational symmetry.  $F_a$  values were calculated by AFRO (<http://www.bfsc.leidenuniv.nl/software/crank/>) and passed into CRUNCH2 (De Graaff *et al.*, 2001), which found six selenium sites. These sites confirmed the translational symmetry seen in the Patterson map. The positions, occupancies, and temperature factors of the sites were refined and phased using BP3 (Pannu *et al.*, 2003). Solvent flattening including the noncrystallographic operator was performed in dm (Cowtan, 1994). For autobuilding and iterative refinement of the model, ARP/wARP (Perrakis *et al.*, 1999) and REFMAC (Murshudov *et al.*, 1999), with the maximum likelihood function incorporating Hendrickson-Lattman coefficients (Pannu *et al.*, 1998) from BP3, were used. ARP/wARP was able to build a model containing 495 residues. An anomalous difference Fourier map found six additional anomalous scatterers related by the translational noncrystallographic symmetry. Manual rebuilding of the model and addition of the unbuilt residues were done with Coot (Emsley and Cowtan, 2004). Refinement to 1.55 Å resolution was done using REFMAC including noncrystallographic symmetry restraints for residues 1 - 277, and water molecules were added using ARP/wARP and Coot. Illustrations were prepared using PyMOL (Figures 1, 2, 5B, 5D, and 6; DeLano, 2002), and an Accelrys DS visualizer (Figure 5C; <http://www.accelrys.com/>).

## RESULTS AND DISCUSSION

### Structure determination

The structure of UVDE was determined to a resolution of 1.55 Å by three-wavelength multiple anomalous dispersion (MAD) using the anomalous signal from selenium atoms. The protein crystallized in space group P1, with unit cell dimensions of 47.34 × 48.70 × 68.76 Å and angles of  $\alpha = 106.1^\circ$ ,  $\beta = 94.4^\circ$ , and  $\gamma = 114.2^\circ$ . The crystals contained two molecules in the asymmetric unit, had a solvent content of 38.2 %, and a Wilson temperature factor of 17.3 Å<sup>2</sup>. The resulting map showed good electron density for the residues -2 to 277 (monomer A) and 1 to 277 (monomer B). The His tag and most of the linker residues and the C-terminal residues 278 - 280 were disordered. There are only small differences between monomers A and B, which are related by a noncrystallographic translational symmetry vector of  $\sim 0 \frac{1}{2} \frac{1}{2}$ . The two molecules in the asymmetric unit are very similar: the root-mean-square deviation (rmsd) between the atomic positions of corresponding main-chain atoms of the two monomers is 0.23 Å. The final model has good stereochemistry and R factors (Table 1) and it contains 4433 protein atoms, 6 anomalously scattering metal ions, 2 phosphate ions, and 356 water molecules.

**Table 1.** Crystallographic data and refinement statistics.

	Peak set	Inflection point	High remote
<i>A. Data collection</i>			
Beam-line	ESRF BM14	ESRF BM14	ESRF BM14
Wavelength (Å)	0.97800	0.97850	0.91840
Detector	MAR225 CCD	MAR225 CCD	MAR225 CCD
Resolution range (Å)	64.6-1.6 (1.69-1.60) <sup>a</sup>	23.6-1.55 (1.63-1.55) 24.0-1.50 (1.58-1.50)	
Multiplicity	3.9 (3.7)	3.9 (3.8)	3.9 (3.8)
Completeness (%)	67.9 (15.8) <sup>b</sup>	73.8 (23.6)	78.9 (31.6)
Rsym <sup>c</sup> (%)	4.0 (33.1)	3.9 (31.5)	4.4 (34.9)
<i>B. Phasing</i>			
Number of Se-sites		6	
FOM Overall		0.433	
FOM 1.73 – 1.55 Å		0.178	
<i>C. Refinement</i>			
Resolution range (Å)		20.0-1.55 (1.63-1.55)	
No. of reflections used in refinement	56303 (1053)		
No. of reflections used for R-free	2860 (66)		
R-factor <sup>d</sup>		0.18 (0.25)	
R-free		0.21 (0.35) <sup>e</sup>	
No. of protein / water atoms	4433 / 356		
Average B-value	protein / solvent (Å <sup>2</sup> ) phosphate / metal ions	18.6 / 28.4 40.0 / 18.5, 68.7, 74.4	
Ramachandran statistics <sup>f</sup> (%)	91.1 / 8.0 / 0.4 / 0.4		
R.m.s. deviations <sup>f</sup> (bonds, Å / angles, °)	0.015 / 1.5		

<sup>a</sup> Values in parentheses are for the highest resolution bin, where applicable.

<sup>b</sup> Data in the higher resolution shell are less complete because of data collection on a square detector.

<sup>c</sup>  $R_{sym} = \sum_h \sum_i |I_{hi} - \langle I_h \rangle| / \sum_h \sum_i |I_{hi}|$ , where  $I_{hi}$  is the intensity of the  $i^{\text{th}}$  measurement of the same reflection and  $\langle I_h \rangle$  is the mean observed intensity for that reflection.

<sup>d</sup>  $R = \sum ||F_{obs}(hkl)| - |F_{calc}(hkl)| | / \sum |F_{obs}(hkl)|$ .

<sup>e</sup> According to the program PROCHECK (Laskowski, 1993). The percentages are indicated of residues in the most favored, additionally allowed, generously allowed and disallowed regions of the Ramachandran plot, respectively.

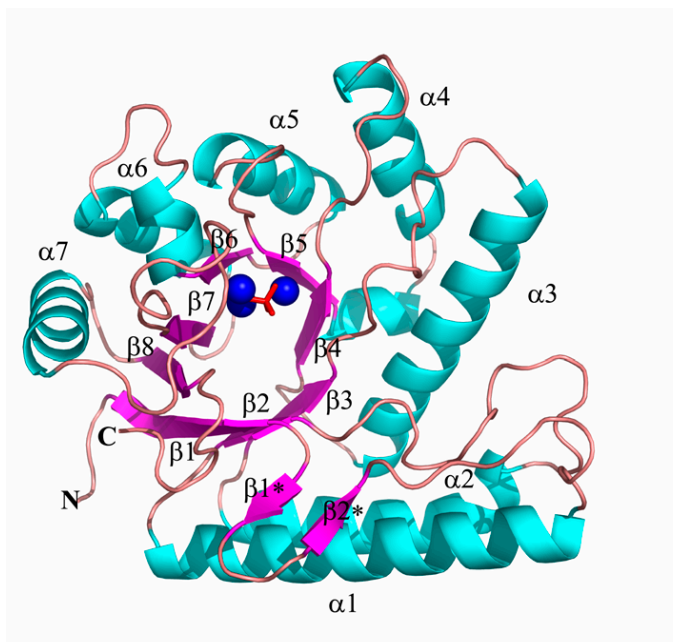
<sup>f</sup> Estimates provided by the program REFMAC (Murshudov *et al.*, 1999).

<sup>g</sup> Bin free R value set count is 66 reflections.

## Structure overview and the active site

The crystal structure shows UVDE to be a single-domain TIM barrel lacking the  $\alpha 8$  helix of the prototypical TIM-barrel fold (Figure 1). Apart from the eight central  $\beta$  strands forming the barrel, two additional  $\beta$  strands ( $\beta 1^*$  and  $\beta 2^*$ ) are present in the N-terminal part of the protein, before the  $\alpha 1$  helix. Unexplained density at the tip of the side chain of Lys229 suggested this residue

may be modified, but the biochemical characterization and potential role of this modification are beyond the scope of this paper.

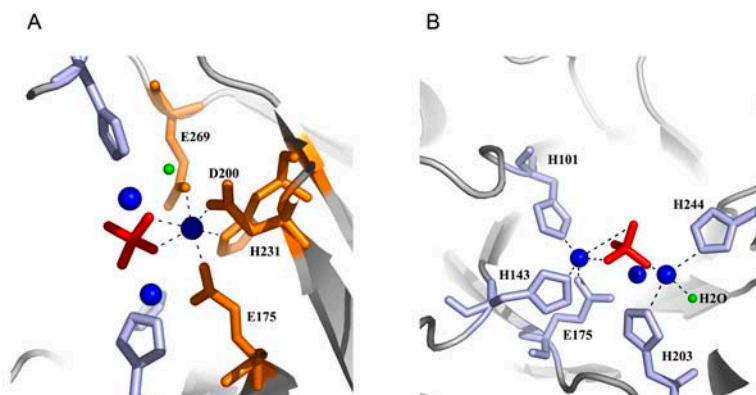


**Figure 1.** UVDE secondary structure.

View of the UVDE overall fold and topology. The  $\alpha$  helices and  $\beta$  strands are labelled according to the canonical TIM-barrel fold and colored light blue and purple respectively. The metal ions are colored blue with the coordinating phosphate in red.

The refined structure of UVDE has three anomalously scattering metal ions, located closely to the C terminus. This, together with the close proximity of the protein's N and C termini, classifies UVDE as a member of the TIM-barrel family of divalent metal-dependent enzymes. The metal ions are well ordered, with a mean temperature factor of 18.5, 68.7, and 74.4  $\text{\AA}^2$ . Because the occupancies of all metal ions were set to one, the high mean temperature factors for two of them probably represent low occupancy rather than high mobility. One of the metal ions is octahedrally coordinated by the side chains of residues Glu175, Glu269, His231, and Asp200 and two oxygen atoms from a phosphate ion (Figure 2A). Another metal ion shows distorted bipyramidal coordination by His101, His143, and Glu175 and two oxygen atoms from the phosphate ion (Figure 2B), while the third metal has an irregular four-fold coordination by one oxygen atom from the phosphate, His244, His203, and one water molecule (Figure 2B).

As a result of the low occupancy for two of the three metal ions, a fluorescence scan performed at the beamline was unable to determine the nature of the metal ions. Because we did not include any metal ions in the crystallization buffer, UVDE must have picked them up



**Figure 2.** Metal coordination.

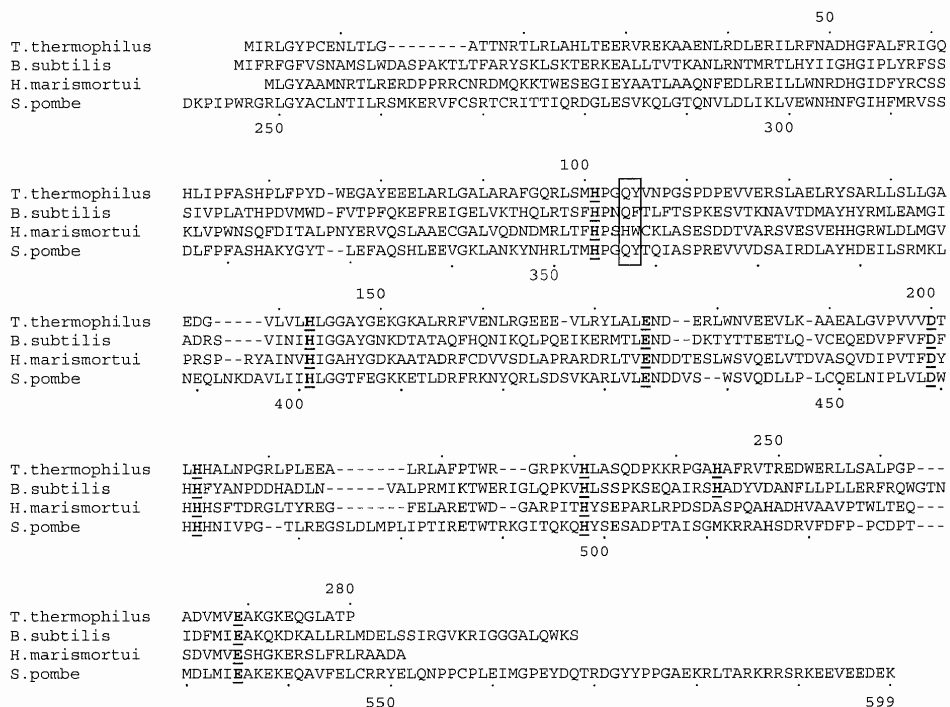
A. The octahedrally coordinated metal ion is colored blue. The four coordinating residues, H231, D200, E269 and E175 are shown in ball-and-stick representation and colored orange. The phosphate coordinated by two oxygen atoms is colored red.

B. The distorted bipyrimidal coordination of the second metal ion by His101, His143, Glu175 and two oxygen atoms from the phosphate ion, while the third metal ion has an irregular four fold coordination by one oxygen atom from the phosphate, His244, His203 and one water molecule. The residues involved in the coordination are colored in light blue (H101, H143, H244 and H203). The phosphate is colored red, the water molecule is in green and the metal ions are colored blue.

from its heterologous expression host *Escherichia coli* or during the purification procedure, which included the use of a nickel column. For incision of DNA containing CPD and 6-4PP UV lesions, however, the additional presence of 1 mM  $Mn^{2+}$  was required (see below). Inclusion of 10 mM  $Mg^{2+}$  instead of  $Mn^{2+}$  could only marginally activate the enzyme: CPD-damaged DNA was not incised, whereas the 6-4PP-containing substrate was incised with extremely low efficiency. Apparently  $Mn^{2+}$  is a required cofactor, and possibly one or more metal sites picked up nickel during the purification procedure, which has to be exchanged with manganese for enzyme activity. However, it remains to be determined how many ordered  $Mn^{2+}$  ions are required for full activity of UVDE. These observations agree with the structurally related enzymes Endo IV (Hosfield *et al.*, 1999) and xylose isomerase (Carrell *et al.*, 1989), which need a cluster of three or two divalent ions for catalytic activity, respectively.

Except for H244, all the metal-coordinating residues are fully conserved in all known UVDE homologs (Figure 3), and we tested the importance of some of these residues in mutational studies. Residue Glu175 was mutated into an alanine (E175A) and the incision efficiency of this point mutant was compared to that of the wild-type protein using 5'-labelled substrates containing either a CPD or a 6-4PP. Incubation of these substrates with wild-type UVDE resulted in incision efficiencies of 95 % and 90 %, respectively (Figures 4A and 4B, lane 2, both panels). Incubation of the same substrates with the E175A mutant did not result in any detectable incision on the

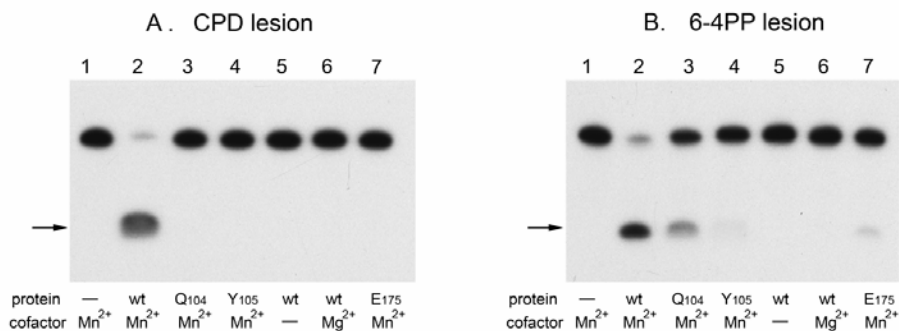
CPD substrate and a very low incision on the 6-4PP-containing substrate (Figures 4A and 4B, lane 7, both panels). Having a severe kink of  $\sim 44^\circ$ , the DNA duplex is more distorted in the 6-4PP than in the CPD substrate (Kim and Choi, 1995). This significant distortion might facilitate the recognition of the 6-4PP damage, which could result in the observed residual activity of the catalytically impaired UVDE E175A mutant.



**Figure 3.** Alignment of UVDE homologues.

The amino acid sequence of UVDE from *T. thermophilus* (Henne *et al.*, 2004) is aligned with a homologue from another eubacterium, *B. subtilis* (Kunst *et al.*, 1997), a homologue from an archaeobacterium *H. marismortui* (Baligal *et al.*, 2004) and a homologue from a eukaryote, *S. pombe* (Takao *et al.*, 1996). Note that the *S. pombe* protein has an additional N-terminal extension of 240 amino acids. The metal-coordinating residues (H101, H143, E175, D200, H203, H231, H244 and E269) are in bold and underlined. The Gln104 and Tyr105 residues proposed to intercalate the DNA are boxed.

Mutating homologous metal-coordinating residues in *S. pombe* UVDE (Glu434, Asp459, and His498; see Figure 3) also resulted in proteins with impaired incision. No detectable incision was observed with mutants E434A and D459A, either on the CPD or on the 6-4PP, showing the crucial role of these residues for the *S. pombe* catalytic site. Mutant H498A showed a reduced activity on the two substrates, indicating that in the absence of this residue the active site metals can still bind, albeit with lower affinity. Taken together, these data suggest that the UVDE active



**Figure 4.** DNA incision by UVDE. Terminally labeled 30 bp DNA substrates with a CPD.

A or 6-4PP lesion B were incubated with (mutant) UVDE protein in the presence or absence of the indicated metal ions. The incision product is indicated with an arrow. The lanes marked Q104, Y105 and E175 contain the corresponding alanine substitutions at these positions.

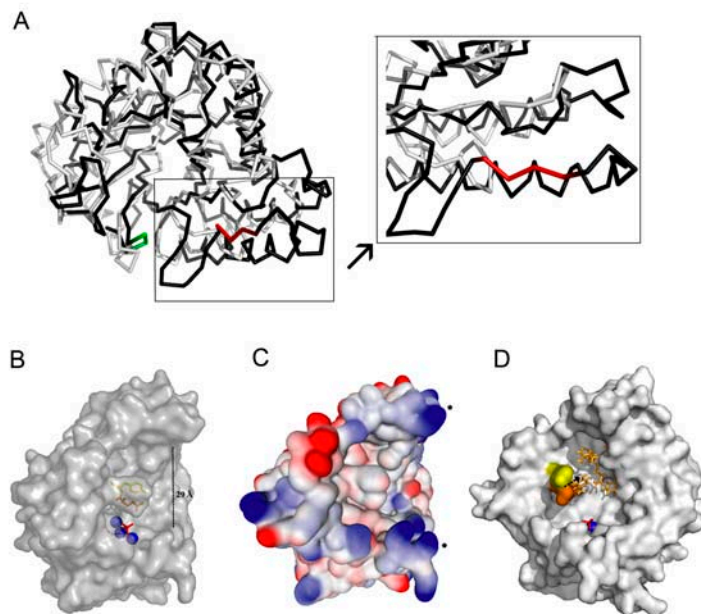
site metal ions are likely to be directly involved in phosphodiester cleavage, as observed in other DNA repair enzymes such as Endo IV. Indeed, the three metal sites that were observed superimpose on the three zinc sites of the Endo IV structure. The positively charged metal ions can act as a Lewis acid to stabilize a water-derived hydroxide attacking the phosphodiester backbone of DNA and counteracting the developing negative charge on the DNA during the cleavage reaction.

### Comparison with endo IV and potential interactions with damaged DNA

Superposition of UVDE with the DNA-repair enzyme Endo IV (Hosfield *et al.*, 1999) from the BER pathway shows that the two enzymes share major structural features, including the TIM-barrel fold (Figure 5A) and a wide groove (29 Å) that houses the active site at the bottom (Figure 5B). An extensive positive charge on both sides of the solvent-accessible groove of UVDE (Figure 5C) can be seen, which is suited to bind a DNA duplex.

However, there are important differences, which reflect the functional divergence between the two enzymes: Endo IV only nicks at abasic sites, whereas UVDE also recognizes other types of damage. Moreover, incision of the *T. thermophilus* UVDE on the abasic site lesion is only 20 % (Figure 6), which is lower compared to the reported efficiency of *S. pombe* UVDE (Kanno *et al.*, 1999). The difference in substrate specificity of the *T. thermophilus* homolog in comparison with UVDE proteins from other species will be addressed elsewhere (unpublished data).

In the structure of Endo IV in complex with a DNA duplex nicked at the abasic site, the DNA helix has a ~ 90° kink and both the abasic phosphoribose and the base opposing this damage are flipped out of the double helix (Hosfield *et al.*, 1999). The nicked, abasic DNA of the Endo IV-DNA complex could be fitted analogously into UVDE: the deep groove of UVDE can comfortably harbour the kinked duplex. The active sites of UVDE and Endo IV are both located at the bottom of the groove.



**Figure 5.** DNA binding site.

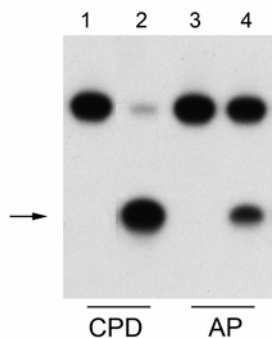
A.  $\alpha$  superpositions of UVDE and Endo IV showing that both enzymes share major structural features. UVDE is shown in black and Endo IV in gray. Residues 18 - 21 in the NRTL-strand of UVDE causing a clash between the DNA in UVDE are colored red. The residues Lys273 and Glu274, which need to make a small rearrangement for fitting of the flipped-out base, are colored green. The superposition was done using the program Theseus (Theobald and Wuttke, 2006)

B. Surface representation of UVDE showing the (semi)conserved residues Tyr105 and Gln104 in ball-and-stick representation and colored in yellow and orange respectively. The metal ions are colored blue and the coordinating phosphate is in red.

C. Electron surface potential of UVDE. Positive charges are marked in blue, and negative charges are in red. The positively charged rims of the groove (marked with \*) suggest the DNA binding site.

D. Surface representation of UVDE showing the cavity allowing Q104 and Y105 movement. The surface of UVDE is rotated 90° clockwise compared to B and C. The surfaces of residues Q104 and Y105 are colored yellow and orange, respectively. The possible movement of Q104 and Y105 is indicated with a black arrow. The residues of Endo IV (Phe32, Asn35, Gln36, Arg37 and Tyr72) that show that this cavity is not present in Endo IV are colored orange and are shown in ball-and-stick representation.

At the bottom of the proposed DNA binding groove of UVDE is a loop with the conserved sequence GQY, in which Gln104 and Tyr105 point straight into the solvent (Figure 5B). Endo IV has the same loop, albeit with different, Endo IV-specific residues. In Endo IV, the side-chain residues Tyr72 and, to a lesser extent, Leu73, project into the kink of the DNA duplex (Hosfield *et al.*, 1999). We propose that the two loops are functionally equivalent and that Gln104 and



**Figure 6.** Incision of an abasic site DNA lesion by UVDE.

Terminally labelled 30 bp DNA substrates with a CPD or an Abasic site (AP) lesion (as indicated) were incubated with *T. thermophilus* UVDE protein (lanes 2 and 4) in the presence of 1 mM  $Mn^{2+}$ . The incision product is indicated with an arrow.

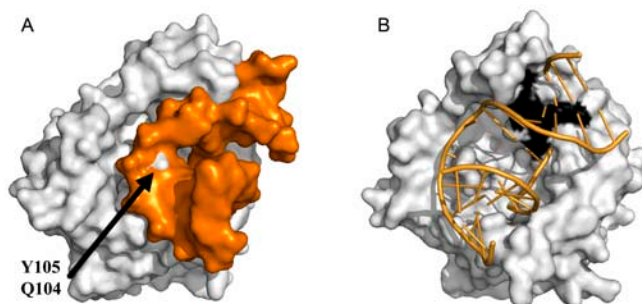
Tyr105 of the UVDE GQY loop stabilize the kink in the DNA duplex at the position of damage in a similar fashion. For Gln104 and Tyr105 to take up a similar position as Tyr72 and Leu73, the loop needs to shift by about 3 Å upon DNA binding, while the side chains may need to adopt a different conformation. There is a cavity in the UVDE structure that allows such a shift and the presence of the conserved Gly103 suggests potential flexibility of this loop (Figure 5D).

In UVDE, the close proximity of Gln104 and Tyr105 to the metal coordination site suggests that these residues might similarly probe the DNA for damage and present the scissile phosphodiester bond at the 5' side of the lesion to the active site. Tyr105 appears to be semiconserved: in UVDE from other species, phenylalanine or tryptophan can also be found (Figure 3), and these residues could also fulfil the proposed intercalating role.

To confirm the significance of Gln104 and Tyr105 for the UVDE function we created point mutants, changing the residues into alanine (Q104A and Y105A). Indeed, the incision assay with Q104A did not reveal any detectable activity on the CPD (Figure 4A). Some residual activity on the 6-4PP substrate (Figure 4B) is observed, which is again an indication that the 6-4PP might be an easier target for damage recognition and processing. Mutant Y105A did not show any incision on the CPD substrate and an extremely low incision on the 6-4PP (Figures 4A and 4B), underlining the crucial role of Tyr105 for enzyme activity. If DNA binds in the same way in UVDE as in Endo IV, there would be a major clash between the DNA and the NRTL strand of UVDE (comprising residues Asn18 – Leu21; Figures 7A and 7B). There is no equivalent of this strand in Endo IV (Figure 5A, zoom-in). An intriguing possibility is that this strand moves toward the GQY loop, to take up a position similar to the Gln36 – Trp39 loop in Endo IV. Such a movement would reposition Asn18 and Arg19 into the kink of the DNA, where they could make similar interactions with the DNA as Gln36 and Arg37 in Endo IV. These residues in Endo IV are vital for stabilizing the kink and flipping out the base opposing the damage (Hosfield *et al.*, 1999).

The conformational similarities between UVDE and Endo IV suggest that also in UVDE, both the damaged base(s) and the base opposing this damage flip out of the double helix. However, is there space for the damaged base to flip out in UVDE? In Endo IV this is not an issue, as it only recognizes DNA with an abasic site. Space for one or more flipped-out bases in

UVDE can only be created by a small rearrangement of two residues at its C terminus: Lys273 and Glu274 (Figure 5A). The presence of Gly272 and Gly276 just before and after these residues suggests they may indeed have the freedom to move. This movement would create a pocket in which the damaged base could stack against Tyr6, while Lys273, Glu274, and possibly also Asn10 could form polar interactions with the damaged base.



**Figure 7.** DNA binding.

A. Model of UVDE bound to DNA. For modeling the DNA fragment of the Endo IV co-crystal was used (Hosfield *et al.*, 1999). The protein surface is colored in light gray and the DNA is presented in orange surface representation.

B. Same as A, but rotated 90° clockwise. The DNA is presented as a cartoon showing the clash between the DNA and the NRTL-strand of UVDE (amino acids N18-L21 are shown in black surface representation).

## CONCLUSION

Our results shed light on the remarkable ability of UVDE to recognize different types of DNA damage. It is much more versatile than its closest homolog Endo IV, which only nicks adjacent to abasic sites. The two enzymes share some characteristic features: a deep DNA binding groove, at the bottom of which are the catalytic site and residues capable of intercalating in the double helix at the site of damage. Furthermore, we observed three metal ions in the same location as the cluster of three Zn<sup>2+</sup> ions in Endo IV's active site. The intrinsic metals of UVDE, however, seem not to be able to perform an enzymatic activity, but the addition of manganese appears to be required. Superposition of UVDE with Endo IV predicts that the rearrangements in UVDE upon DNA binding are likely to be more substantial than those observed in Endo IV. This structural flexibility might be part of the explanation of the broader substrate specificity of UVDE.

## REFERENCES

- Aravind L, Walker DR and Koonin EV, Conserved domains in DNA repair proteins and evolution of repair systems, *Nucleic Acids Res.* **27** (1999), pp. 1223-1242
- Avery AM, Kaur B, Taylor JS, Mello JA, Essigmann JM and Doetsch PW, Substrate specificity of ultraviolet DNA endonuclease (UVDE/Uve1p) from *Schizosaccharomyces pombe*, *Nucleic Acids Res.* **27** (1999), pp. 2256-2264
- Baliga NS, Bonneau R, Facciotti MT, Pan M, Glusman G, Deutsch EW, Shannon P, Chiu Y, Weng RS and Gan RR, Genome sequence of *Haloarcula marismortui*: a halophilic archaeon from the Dead Sea, *Genome Res.* **14** (2004), pp. 2221-2234
- Carrell HL, Glusker JP, Burger V, Manfre F, Tritsch D and Biellmann JF, X-ray analysis of D-xylose isomerase at 1.9 Å: native enzyme in complex with substrate and with a mechanism-designed inactivator, *Proc. Natl. Acad. Sci. USA.* **86** (1989), pp. 4440-4444
- CCP4 (Collaborative Computational Project, Number 4), The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D Biol. Crystallogr.* **50** (1994), pp. 760-763
- Cowtan K, dm: An automated procedure for phase improvement by density modification, *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* **31** (1994), pp. 34-38
- De Graaff RAG, Hilge M, van der Plas JL and Abrahams JP, Matrix methods for solving protein substructures of chlorine and sulfur from anomalous data, *Acta Crystallogr. D Biol. Crystallogr.* **57** (2001), pp. 1857-1862
- DeLano WL, The PyMOL Molecular Graphics System (<http://www.pymol.org>) (2002)
- Doetsch PW, Beljanski V and Song B, The ultraviolet damage endonuclease (UVDE) protein and alternative excision repair: a highly diverse system for damage recognition and processing. In: V. Beljanski and B. Song, Editors, *DNA Damage Recognition*, Taylor & Francis Press, New York (2006), pp. 211-223
- Emsley P and Cowtan K, Coot: model-building tools for molecular graphics, *Acta Crystallogr. D Biol. Crystallogr.* **60** (2004), pp. 2126-2132
- Evans PR, Data reduction. In: E. Dodson, M. Moore, A. Ralph and S. Bailey, Editors, *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, Daresbury Laboratory, Warrington, UK (1993), pp. 114-122
- Friedberg EC, Walker GC, Siede W, Wood RD, Schultz AR and Ellenberger T, *DNA Repair and Mutagenesis*, ASM Press, Washington, DC (2005)
- Henne A, Bruggemann H, Raasch C, Wiezer A, Hartsch T, Liesegang H, Johann A, Lienard T, Gohl O and Martinez-Arias R, *et al.*, The genome sequence of the extreme thermophile *Thermus thermophilus*, *Nat. Biotechnol.* **22** (2004), pp. 547-553
- Hosfield DJ, Guan Y, Haas BJ, Cunningham RP and Tainer JA, Structure of the DNA repair enzyme Endo IV and its DNA complex: double-nucleotide flipping at abasic sites and three-metal-ion catalysis, *Cell* **98** (1999), pp. 397-408
- Iwai S, Chemical synthesis of oligonucleotides containing damaged bases for biological studies, *Nucleosides Nucleotides Nucleic Acids* **25** (2006), pp. 561-582
- Kanno S, Iwai S, Takao M and Yasui A, Repair of apurinic/apyrimidinic sites by UV damage endonuclease: a repair protein for UV and oxidative damage, *Nucleic Acids Res.* **27** (1999), pp. 3096-3103
- Kaur B, Avery AM and Doetsch PW, Expression, purification, and characterization of ultraviolet DNA endonuclease from *Schizosaccharomyces pombe*, *Biochemistry* **37** (1998), pp. 11599-11604
- Kaur B, Fraser JL, Freyer GA, Davey S and Doetsch PW, A Uve1p-mediated mismatch repair pathway in *Schizosaccharomyces pombe*, *Mol. Cell. Biol.* **19** (1999), pp. 4703-4710

- Kim JK and Choi BS, The solution structure of DNA duplex-decamer containing the (6-4) photoproduct of thymidylyl (3'→5') thymidine by NMR and relaxation matrix refinement, *Eur. J. Biochem.* **228** (1995), pp. 849-854
- Kunst F, Ogasawara N, Moszer I, Albertini AM, Alloni G, Azevedo V, Bertero MG, Bessieres P, Bolotin A and Borchert S, *et al.*, The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*, *Nature* **390** (1997), pp. 249-256
- Laskowski RA, PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Cryst.* **26** (1993), pp. 283-291
- Leslie AG, Integration of macromolecular diffraction data, *Acta Crystallogr. D Biol. Crystallogr.* **55** (1999), pp. 1696-1702
- Murshudov GN, Vagin AA, Lebedev A, Wilson KS and Dodson EJ, Efficient anisotropic refinement of macromolecular structures using FFT, *Acta Crystallogr. D Biol. Crystallogr.* **55** (1999), pp. 247-255
- Pannu NS, Murshudov GN, Dodson EJ and Read RJ, Incorporation of prior phase information strengthens maximum likelihood structure refinement, *Acta Crystallogr. D Biol. Crystallogr.* **54** (1998), pp. 1285-1294
- Pannu NS, McCoy AJ and Read RJ, Application of the complex multivariate normal distribution to crystallographic methods with insights into multiple isomorphous replacement phasing, *Acta Crystallogr. D Biol. Crystallogr.* **59** (2003), pp. 1801-1808
- Perrakis A, Morris R and Lamzin V.S, Automated protein model building combined with iterative structure refinement, *Nat. Struct. Biol.* **6** (1999), pp. 458-463
- Sohi M, Alexandrovich A, Moolenaar GF, Visse R, Goosen N, Vernede X, Fontecilla-Camps J.C, Champness J and Sanderson M.R, Crystal structure of *Escherichia coli* UvrB C-terminal domain, and a model for UvrB-UvrC interaction, *FEBS Lett.* **465** (2000), pp. 161-164
- Studier FW, Rosenberg AH, Dunn JJ and Dubendorf JW, Use of T7 RNA polymerase to direct expression of cloned genes, *Methods Enzymol.* **185** (1990), pp. 60-89
- Takao M, Yonemasu R, Yamamoto K and Yasui A, Characterization of a UV endonuclease gene from the fission yeast *Schizosaccharomyces pombe* and its bacterial homolog, *Nucleic Acids Res.* **24** (1996), pp. 1267-1271.
- Theobald DL and Wuttke DS, Theseus: maximum likelihood superpositioning and analysis of macromolecular structures, *Bioinformatics* **22** (2006), pp. 2171-2172
- Truglio JJ, Croteau DL, Van Houten B and Kisker C, Prokaryotic nucleotide excision repair: the UvrABC system, *Chem. Rev.* **106** (2006), pp. 233-252
- Verhoeven EE, van Kesteren M, Turner JJ, van der Marel GA, van Boom JH, Moolenaar GF and Goosen N, The C-terminal region of *Escherichia coli* UvrC contributes to the flexibility of the UvrABC nucleotide excision repair system, *Nucleic Acids Res.* **30** (2002), pp. 2492-2500



# Involvement of a carboxylated lysine in UV damage endonuclease

Chapter

3

Elisabeth M. Meulenbroek<sup>1,2</sup>, Keti Paspaleva<sup>2</sup>, Ellen A.J. Thomassen<sup>1</sup>, Jan Pieter Abrahams<sup>1</sup>, Nora Goosen<sup>2</sup>, Navraj S. Pannu<sup>1</sup>

---

<sup>1</sup>Biophysical Structural Chemistry, Leiden Institute of Chemistry, P.O. Box 9502, 2300 RA Leiden, The Netherlands

<sup>2</sup>Laboratory of Molecular Genetics, Leiden Institute of Chemistry, P.O. Box 9502, 2300 RA Leiden, The Netherlands

## SUMMARY

UV damage endonuclease is a DNA repair enzyme that can both recognise damages such as UV lesions and introduce a nick directly 5' to them. Recently, the crystal structure of the enzyme from *Thermus thermophilus* was solved. In the electron density map of this structure unexplained density near the active site was observed at the tip of Lys229. Based on this finding, it was proposed that Lys229 is post-translationally modified. In this paper, we give evidence that this modification is a carboxyl group. By combining activity assays and X-ray crystallography on several point mutants, we show that the carboxyl group assists in metal-binding required for catalysis by donating negative charge to the metal-coordinating residue His231. Moreover, functional and structural analysis of the K229R mutant reveals that if His231 shifts away, resulting in an increased activity on both damaged and undamaged DNA. Taken together, the results show that *T. thermophilus* UVDE is carboxylated and the modified lysine is required for proper catalysis and preventing increased incision of undamaged DNA.

## INTRODUCTION

Repairing damage in DNA is essential for maintaining genomic integrity. Therefore, several protein systems have evolved in order to remove DNA lesions. One of them is the ultraviolet damage endonuclease (UVDE) repair system, initially found in the yeast *Schizosaccharomyces pombe*, and described as an alternative repair system for UV-induced lesions (Sidik *et al.*, 1992). The UVDE enzyme was shown to introduce a nick 5' to both of the main UV-induced lesions: cyclobutane pyrimidine dimers (CPD) and 6-4 photoproducts (6-4PP) (Bowman *et al.*, 1994). Later studies showed, however, that UVDE from *S. pombe* has much broader substrate specificity than originally thought - recognising and nicking DNA lesions significantly different from UV induced damage such as abasic sites and small loops (Avery *et al.*, 1999). UVDE enzymes have been found in both eukaryotes (e.g. *S. pombe*, *Neurospora crassa*) and in prokaryotes (e.g. *Thermos thermophilus*, *Bacillus subtilis*).

The crystal structure of the *T. thermophilus* UVDE has recently been solved to 1.55 Å resolution (Paspaleva *et al.*, 2007). This structure shows UVDE to be a single domain TIM-barrel with extensive positive charges positioned on both sides of a 29 Å groove, which was proposed to be the DNA-binding site. At the bottom of the groove, a cluster of three metal ions was found. Furthermore, a protein pocket was identified in UVDE near Tyr6 and Asn10, into which UVDE might flip damaged base(s).

Unexplained electron density at the tip of Lys229 was observed and it was suggested that Lys229 is post-translationally modified. Here, we investigate the identity of this modification by X-ray crystallography and conclude that it is a carboxyl group. Carboxylated lysines have been observed in several proteins (Abendroth *et al.*, 2002; Cha and Mobashery 2007; Golema-Kotra *et al.*, 2003; Golemi *et al.*, 2001; Li *et al.*, 2005; Thoden *et al.*, 2001, 2003). Most often, the carboxyl group has a structural role (e.g. bridging two active site zinc atoms), but a more direct role in reaction mechanisms has also been reported (Cha and Mobashery 2007). In addition to identifying this modification, we study, by means of X-ray crystallography and activity assays on several point mutants, why the protein uses a carboxylated lysine instead of a standard amino acid.

## EXPERIMENTAL PROCEDURES

### Proteins

The *T. thermophilus* UVDE protein used in this study was expressed and purified as described before (Paspaleva *et al.*, 2007). Mutants K229A, K229L and K229R were constructed by site-directed mutagenesis using PCR and purified in the same way as the wild type enzyme.

### DNA substrates

DNA substrates used in all activity assays are 30 bp DNA containing either a CPD or a 6-4PP in the following sequence: 5'-CTCGTCAGCATCTTCATCATAACAGTCAGTG-3' with **TT** representing the position of the UV lesion. In case of the abasic site the same DNA sequence has been used: 5'-CTCGTCAGCATC**X**TTCATCATAACAGTCAGTG-3', with **X** representing the position of the abasic site. The oligonucleotides containing CPD or 6-4PP lesions were synthesised as described (Iwai, 2006). The 30bp substrate containing an abasic site lesion has been obtained commercially (Eurogentec, Belgium).

### Incision assay

The DNA substrates were labelled at the 5' side of the top strand using polynucleotide kinase as described (Verhoeven *et al.*, 2002). The DNA substrates (0.2 nM) were incubated with 5 nM UVDE in 20  $\mu$ l reaction mix (20 mM HEPES pH 6.5, 100 mM NaCl, 1 mM MnCl<sub>2</sub>). After 15 minutes incubation at 55°C the reaction was terminated by adding 3  $\mu$ l EDTA/SDS (0.33 M EDTA, 3.3 % SDS) and 2.4  $\mu$ l glycogen (4  $\mu$ g/ $\mu$ l) followed by ethanol precipitation. The incision products were loaded on a 15 % denaturing polyacrylamide gel and visualised by irradiation of a photographic film after which they were quantified.

### Incision of supercoiled plasmid DNA

Supercoiled plasmids pUC18 (2686 bp; 5 ng/ $\mu$ l; UV-irradiated at 300 J/m<sup>2</sup> or not UV-irradiated) and pNP228 (4686 bp; 5 ng/ $\mu$ l) were incubated with 10 nM UVDE (unless stated otherwise) in a 10  $\mu$ l reaction mix (20 mM HEPES pH 6.5, 100 mM NaCl, 1 mM MnCl<sub>2</sub>). After 15 min incubation at 55°C the reactions were terminated by addition of 3  $\mu$ l Ficoll/dyes/SDS/EDTA (0.05 M EDTA, 3 % SDS). Samples were loaded on a 0.7 % agarose gel and visualised by staining with ethidiumbromide.

### Filter Binding Assay

The filter binding assays were conducted in 10  $\mu$ l samples containing 0.1  $\mu$ M UVDE and 8 nM of <sup>32</sup>P-labelled DNA in a reaction buffer containing 20 mM Tris pH 6.5, 100 mM NaCl, 1 mM MnCl<sub>2</sub>. Samples were incubated for 10 min at 55°C. At the end of the incubation time 0.5 mL reaction buffer (preheated at 55°C) was added. The mixture was poured over a nitrocellulose filter and the incubation vial was rinsed with 0.5 ml of the preheated reaction buffer. Finally the filters were washed with 0.5 ml incubation buffer. Each sample was corrected for the amount of DNA retained on a filter in the absence of protein. Binding is expressed as the percentage of the input DNA retained on the filter.

### Mass-spectrometry

Tth-UVDE was precipitated in 10 % trichloroacetic acid (TCA) and subsequently washed with acetone. After drying, it was resuspended in 8 M urea and 0.4 M NH<sub>4</sub>HCO<sub>3</sub>. After 15

min incubation with dithiothreitol at 50°C and addition of iodoacetamide, trypsin digestion was carried out for 24 h at 37°C. The trypsinated sample was loaded on a LTQ-orbitrap mass-spectrometer and MS and MSMS were run.

### **Crystallisation**

The purified protein was dialysed against 1x PBS (25 mM Phosphate, 150 mM NaCl, pH 7.4) and concentrated to 3-5 mg/ml by centrifugation using an Ultrafree Filter Device (Millipore). Crystals were grown by sitting-drop vapour diffusion at 22°C using as precipitant: 0.1 M sodium acetate buffer pH 4.4, 2 M sodium formate and 1 mM MnCl<sub>2</sub> (UVDE pH 4.4; 1 µl protein and 1 µl precipitant) or 0.5 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 M sodium acetate buffer pH 6.0, 1 M Li<sub>2</sub>SO<sub>4</sub> (UVDE K229L and UVDE E175A; 1 µl protein and 0.5 µl precipitant) or 0.1 M sodium acetate buffer pH 5.6, 1 M sodium formate (UVDE K229R; 1 µl protein and 1 µl precipitant). The crystals grew like diamonds (40 to 100 µm) in 1-2 days. Crystals were transferred to precipitant solution with 10 % glycerol prior to data collection.

### **Data collection and processing**

Datasets were collected on beam-line ID14-2 at the European Synchrotron Radiation Facility at a wavelength of 0.933 Å. The crystals were flash-frozen and kept at 100 K during data collection. Data collection strategy was determined with the program BEST (Popov and Bourenkov, 2003). Reflections were indexed and integrated with iMosflm (Leslie, 1999). Scaling and merging was performed with SCALA from the CCP4 (CCP4 1994) suite. For data statistics, see Table 1.

### **Structure solution and refinement**

Structure solution was performed by molecular replacement with as search model the structure of Tth-UVDE wild type (pdb-code: 2j6v) with an unmodified lysine using MOLREP (Vagin and Teplyakov, 1997) from the CCP4 suite, though for K229R, the structure of UVDE pH 4.4 was used as input directly into refinement. Manual adjustments to the model were done with COOT (Emsley and Cowtan, 2004). Refinement was done using REFMAC (Murshudov *et al.*, 1997) and water molecules were added using ARP/wARP (Perrakis *et al.*, 1999) and COOT. TLS refinement was used in refinement for structures UVDE pH 4.4, UVDE K229L and UVDE E175A. Illustrations were prepared using CCP4mg (Potterton *et al.*, 2004).

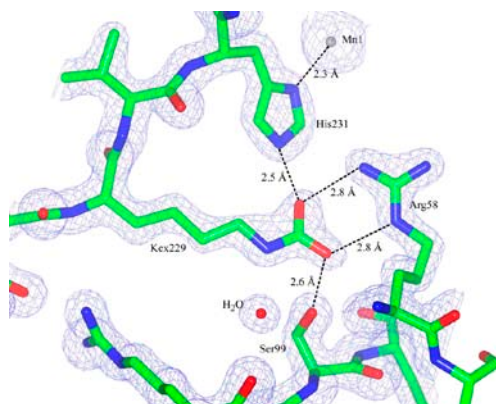
## **RESULTS**

### **Identity of the modification**

The unexplained electron density observed at the tip of Lys229 is situated between Arg58, Ser99 and His231. His231 was shown to coordinate one of the metal ions (Mn1). By comparing the metal ion cluster of Tth-UVDE to that of endonuclease IV (Hosfield *et al.*, 1999), Mn1 is

probably one of the two metal ions that together activate the water molecule (bridging Mn1 and Mn2) that serves as a nucleophile to attack the DNA phosphodiester backbone.

Several options are possible for the unexplained electron density at the tip of Lys229 among which is an acetyl group or a carboxyl group. Considering the potential hydrogen bonding to the neighbouring residues (Arg58, Ser99 and His231), a carboxyl group is the more likely option (Figure 1). Moreover, if an acetyl group is modelled into the density, its methyl group's B-factor is 13.94 Å<sup>2</sup> while the N $\zeta$  and C $\epsilon$  of the lysine and the carbonyl of the acetyl group all have B-factors between 20 and 22 Å<sup>2</sup>. If a carboxyl group is modelled into the density, all atoms of the carboxyl group and the N $\zeta$  and C $\epsilon$  of the lysine have B-factors between 18 and 21 Å<sup>2</sup>.



**Figure 1.** Detail of the original structure of Tth-UVDE wild type with electron density map (contoured at 1.5  $\sigma$ ) showing a carboxyl group modelled into the additional density at the tip of Lys229 (Kcx229).

Carbons are depicted in green, oxygens in red, nitrogens in blue and the metal ion in grey. Distances between the oxygen atoms of the carboxyl group and neighbouring residues, and the distance of His231 to Mn1 are indicated.

To begin the investigation on the modification's identity we performed mass-spectrometry on trypsinated UVDE. In the resulting mass-spectrum, however, only a peptide fragment with an unmodified Lys229 was found. Considering the preparation of the sample (TCA precipitation and trypsin digestion), an acetyl group probably would have been observed if it had been present. In contrast, a carboxyl group on a lysine residue can be expected to fall off during the sample preparation, since at low pH the  $\zeta$ -nitrogen of the lysine is protonated, which results in a loss of the carboxyl group as carbon dioxide.

To confirm the above result, we obtained UVDE crystals at low pH (expecting that the carboxylation will not be present, due to the acid liability of the carboxylation). For this, Tth-UVDE was crystallised in 0.1 M acetate buffer pH 4.4, 2 M sodium formate and 1 mM MnCl<sub>2</sub> (UVDE pH 4.4) and crystallised in a different space group (P 6<sub>1</sub>22) with different packing than the wild type structure (P1) (See Table 1 for the crystallographic statistics).

**Table 1.** Crystallographic data and refinement statistics.

	UVDE pH 4.4	UVDE K229L	UVDE K229R	UVDE E175A
<i>a. Data collection</i>				
Beam line	ESRF ID14-2	ESRF ID14-2	ESRF ID14-2	ESRF ID14-2
Detector	MAR225 CCD	MAR225 CCD	MAR225 CCD	MAR225 CCD
Resolution range (Å)	28.78-1.91 Å (2.01-1.91 Å) <sup>a</sup>	37.11-2.30 Å (2.42-2.30 Å)	46.37-3.15 Å (3.32-3.15 Å)	40.46-2.74 Å (2.89 – 2.74 Å)
Multiplicity	6.4 (6.5)	11.3 (11.5)	10.7 (11.0)	4.4 (4.4)
Completeness (%)	99.2 (100)	99.2 (99.8)	95.3 (96.2)	99.3 (99.8)
R <sub>merge</sub> <sup>b</sup> (%)	0.048 (0.282)	0.078 (0.367)	0.134 (0.366)	0.087 (0.364)
R <sub>pim</sub> <sup>c</sup> (%)	0.020 (0.113)	0.025 (0.122)	0.043 (0.117)	0.048 (0.203)
I/sigma	10.2 (2.0)	8.3 (2.0)	5.4 (2.0)	8.0 (2.0)
Space group	P6 <sub>1</sub> 22	P6 <sub>1</sub> 22	P6 <sub>1</sub> 22	P6 <sub>1</sub> 22
N° of molecules in asymmetric unit	1	1	1	1
Unit cell parameters a/b/c (Å)	107.14x107.14x 90.83	113.39x113.39x 89.39	107.11x107.11x 91.12	113.38x113.38x 89.08
Crystallisation condition	0.1 M Ac pH 4.4 2 M Naformate 1 mM MnCl <sub>2</sub>	0.5 M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> 0.1 M Ac pH 6 1 M Li <sub>2</sub> SO <sub>4</sub>	0.1 M Ac pH 5.6 1 M Naformate	0.5 M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> 0.1 M Ac pH 6 1 M Li <sub>2</sub> SO <sub>4</sub>
<i>b. Molecular replacement</i>				
Correlation coefficient	0.613	0.658	NA	0.639
R factor of correct solution/ second peak	0.515/ 0.643	0.442/ 0.625	NA	0.416/ 0.610
<i>c. Refinement</i>				
N° of reflections used in refinement	22835	14615	5122	8773
Cutoff	None	None	None	None
Resolution range (Å)	28.78-1.91 Å (1.96-1.91 Å) <sup>a</sup>	34.28-2.30 Å (2.36-2.30Å)	46.18-3.15 Å (3.23-3.15 Å)	35.03- 2.74 Å (2.81-2.74 Å)
R factor <sup>d</sup>	0.196 (0.286)	0.188 (0.198)	0.195 (0.232)	0.198 (0.295)
R <sub>free</sub> <sup>e</sup>	0.244 (0.363)	0.244 (0.278)	0.249 (0.360)	0.245 (0.363)
Ramachandran statistics <sup>f</sup> (%)	91.1/7.6/0.8/0.4	89.8/8.9/0.4/0.8	91.5/7.7/0.0/0.9	91.9/6.8/0.9/0.4
R.m.s. deviations (bonds, Å/ angle, °) <sup>g</sup>	0.019/ 1.689	0.014/ 1.545	0.006/ 0.945	0.009/ 1.320
Average atomic B-factor for protein/ Mn/ solvent atoms	25.4/ NA/ 33.0	30.6/ 65.2/ 30.6	32.6/ NA/ 32.4	22.8/ 74.3/11.0
Wilson plot B-factor	25.0	35.9	54.4	47.3

<sup>a</sup> Values in parentheses are for the highest resolution bin, where applicable.

<sup>b</sup>  $R_{\text{merge}} = \sum |I - \langle I \rangle| / \sum I$

<sup>c</sup> (Weiss and Hilgenfeld 1997; Diederichs and Karplus 1997)

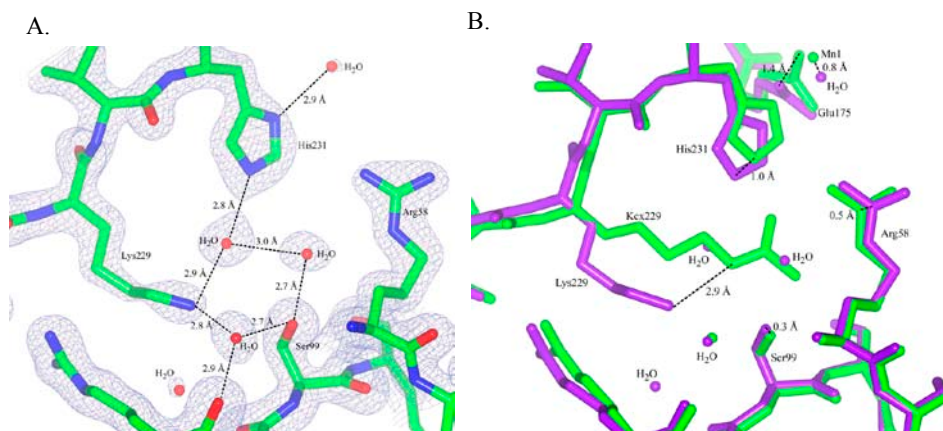
<sup>d</sup>  $R = \sum |F_{\text{obs}}(hkl)| - |F_{\text{calc}}(hkl)| / \sum |F_{\text{obs}}(hkl)|$ .

<sup>e</sup> About 5 % of the reflections were used for the cross-validation set. These reflections were randomly chosen.

<sup>f</sup> According to the program PROCHECK (Laskowski et al. 1993). The percentages are indicated of residues in the most favored, additionally allowed, generously allowed and disallowed regions of the Ramachandran plot, respectively.

<sup>g</sup> Estimates provided by the program REFMAC (Murshudov et al. 1997).

In this structure Lys229 is indeed not modified (Figure 2). Two waters are at the place where the modified lysine originally was. His231 seems to have moved inwards (over 0.96 Å) in order to form hydrogen bonds with one of these waters. This shift seems to have caused a 1.36 Å shift to Glu175. In the structure of UVDE pH 4.4, no metal ions were found, probably because of protonation of the metal-coordinating histidines due to the low pH. A water molecule was found near the site (0.81 Å distance) where originally Mn1 was located.



**Figure 2.** Structure of UVDE pH 4.4.

A. Model and map (contoured at 1.7  $\sigma$ ) of UVDE pH 4.4, showing an unmodified Lys229. Carbons are depicted in green, oxygens in red and nitrogens in blue. Hydrogen network with Ser99, Lys229 and His231 is indicated.

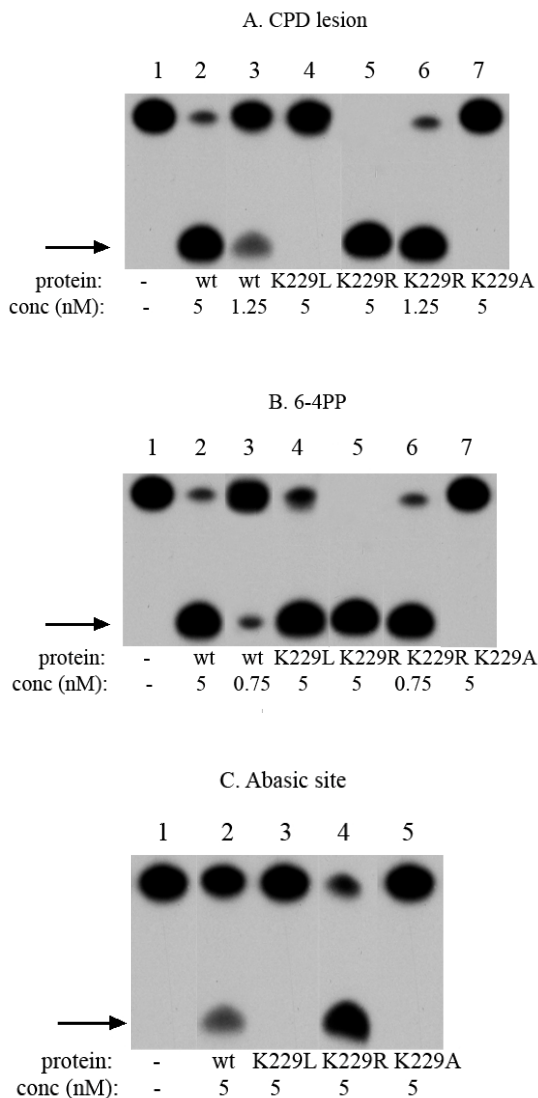
B. Detail of the superposition of the original structure of UVDE (green) and UVDE pH 4.4 (purple). Shifts of residues Arg58, Ser99, Glu175, Lys229 and His231 and of Mn1 compared to the water in UVDE pH 4.4 are indicated at the atoms where the shift was measured.

In order to check that the above result is caused by the low pH of the crystallisation condition and not due to the different space group (P6<sub>1</sub>22) of the previously published structure (P1), we also solved a structure of Tth-UVDE in P6<sub>1</sub>22 that was crystallised at a higher pH (0.5 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 M sodium acetate buffer pH 5.6, 1 M Li<sub>2</sub>SO<sub>4</sub>, 1 mM MnCl<sub>2</sub>; results not shown). In this crystal structure, Lys229 was seen to be modified, showing that the modification can be accommodated in P6<sub>1</sub>22.

### Functional studies of Lys229 mutants

To investigate the role of the carboxylated lysine in the activity of UVDE, we mutated this residue into an alanine (UVDE K229A) and a leucine (UVDE K229L). UVDE K229A showed an extremely reduced catalytic activity on all tested DNA substrates (Figure 3). The K229L mutant showed severely reduced catalytic activity on the CPD (10 %) and the abasic site lesion

(less than 1 %). On the 6-4PP the incision efficiency of this mutant is slightly reduced compared to that of the wild type (Figure 3). These results indicate that Lys229 has an important role in the function of Tth-UVDE.



**Figure 3.** Activity of wild type and mutant UVDE proteins.

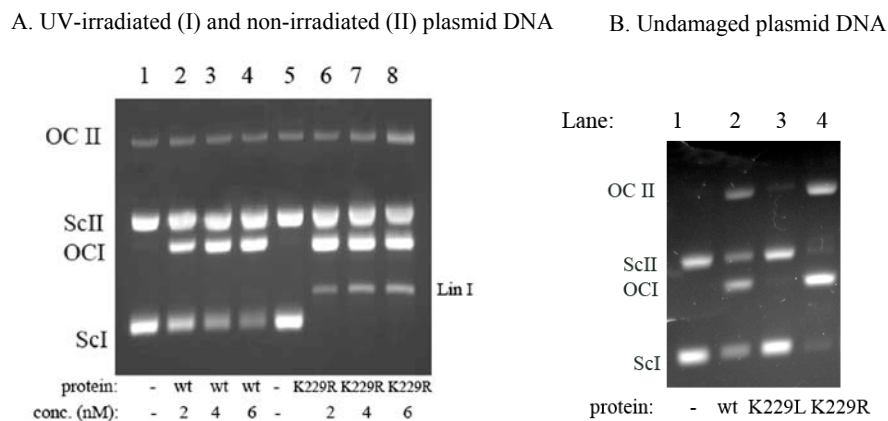
Activity assay with terminally labelled 30 bp DNA substrates containing a CPD (A), 6-4PP lesion (B), abasic site (C). The incision product is indicated with an arrow. Below the lanes is indicated which protein is used and the concentration of protein used (in nM).

To investigate if the observed reduction in the incision efficiency is due to an effect on the formation of protein - DNA complexes, we tested the binding properties of UVDE K229A and UVDE K229L in a filter binding assay. As can be seen in Table 2 the binding capacity of both mutants is similar to the wild type. This indicates that the reduced incision activity is not due to changes in the enzymes ability to bind DNA, but an impairment in catalysis. Moreover, DNA-binding is similar either in absence and presence of  $Mn^{2+}$  (for wild type and mutants) showing that  $Mn^{2+}$  does not influence DNA binding.

To gain a more detailed insight into the function of the carboxylated lysine, a mutant was constructed in which Lys229 was changed to an arginine (UVDE K229R). An arginine has a positive charge like a lysine, but an arginine cannot be carboxylated. The positive side-chain of an arginine can therefore not be turned into a negative group by carboxylation.

Surprisingly, the K229R mutant was found to be active and shows very high activity on CPD (Figure 3A) and 6-4PP (Figure 3B) lesions, even higher than that of wild type. Also the abasic site, which is not an optimal substrate for the Tth-UVDE wild type (20 % incision), is cleaved by the K229R mutant with high efficiency (80 %, Figure 3C).

We also tested the activity of K229R on UV-damaged DNA in a supercoiled incision assay (Figure 4A; plasmid I). In the lanes with UVDE K229R (lane 6 - 8), much more relaxed and even linear plasmid DNA as a result of damage-specific incision can be seen compared to the lanes with UVDE wild type (lanes 2 - 4). This confirms that UVDE K229R is much more active



**Figure 4.** Activity of UVDE wild type and K229R on UV-irradiated and non-irradiated plasmid DNA.

A. A mixture of UV-irradiated supercoiled DNA of pUC18 (I, 2686 bp) and non-irradiated pNP228 (II, 4686 bp) was incubated with different concentrations (as indicated) UVDE wild type and K229R mutant for 15 min. The positions of the supercoiled (sc), open circle (oc) and linear (lin) forms of the plasmids are indicated.

B. A mixture of undamaged supercoiled DNA of pUC18 (I, 2686 bp) and pNP228 (II, 4686 bp) was incubated with 10 nM UVDE wild type and mutant proteins as indicated. The positions of the supercoiled (sc) and open circle (oc) forms of the plasmids are indicated.

on UV damage sites than the wild type protein. UVDE K229R also has incision activity on the undamaged supercoiled plasmid DNA present in the assay (Figure 4A, plasmid II, lane 8), though this activity is very small compared to the activity on UV-damaged supercoiled plasmid DNA. To investigate this activity on undamaged DNA more closely, we performed an assay with only undamaged supercoiled plasmid DNA (Figure 4B). In this assay we could see that, surprisingly, UVDE wild type has some activity on undamaged plasmid DNA (lane 2). The activity of K229R on undamaged DNA, however, was again significantly higher (lane 4).

Filter-binding assays (Table 2) showed that K229R has similar DNA binding properties to wild type UVDE. Thus, the higher activity of this mutant on both damaged and undamaged DNA is not caused by a change to DNA binding, but by an increased efficiency of the incision reaction.

**Table 2.** DNA binding properties of Tth-UVDE wt and mutants tested in a filter binding assay.

DNA lesion	UVDE	% binding no Mn <sup>2+</sup>	% binding 1 mM Mn <sup>2+</sup>
CPD	Wild type	34 ± 2	33 ± 1.7
CPD	K229L	31 ± 0.7	31 ± 0.7
CPD	K229R	31 ± 0.5	32 ± 0.2
6-4PP	Wild type	19 ± 3.7	20 ± 1
6-4PP	K229L	20 ± 0.5	21 ± 1.8
6-4PP	K229R	17 ± 2.8	19 ± 0.7
Abasic site	Wild type	20 ± 2	23 ± 0.7
Abasic site	K229L	21 ± 0.7	21 ± 1.4
Abasic site	K229R	21 ± 2	21 ± 0.7
No damage	Wild type	2.5 ± 0.1	2.8 ± 0.4
No damage	K229L	3.7 ± 0.2	3.8 ± 0.4
No damage	K229R	7.6 ± 0.7	5.8 ± 1

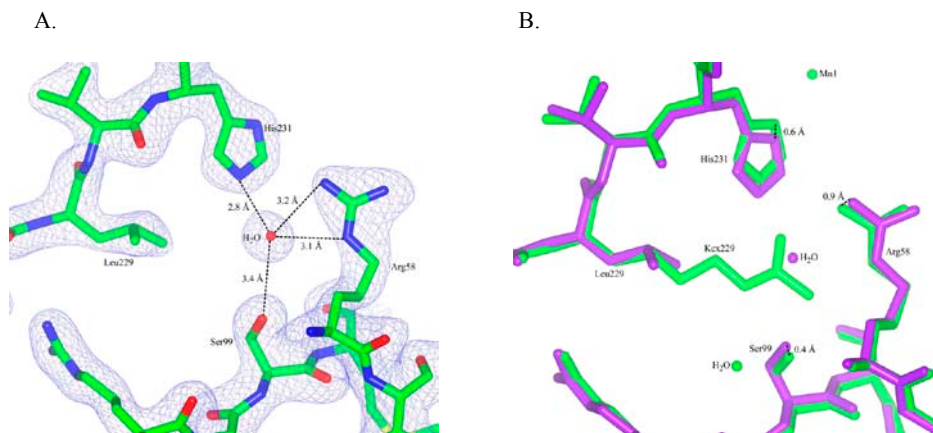
Binding is expressed as the percentage of the input DNA retained on the filter. Binding efficiencies were tested in presence and absence of 1 mM MnCl<sub>2</sub>.

### Structural studies of the Lys229 mutants

To obtain a structural basis for the above results, the mutants K229L, K229R and E175A were crystallised and their structures were solved by molecular replacement (see Table 1 for statistics). In all three mutants, the overall structure is the same as wild type.

In the structure of K229L, a water molecule is present at the place of the carboxylated lysine (Figure 5). Only one of the three metal ions is present: Mn3. This is the metal ion near residues His244 and His203 and it corresponds to Zn3 in endonuclease IV (Hosfield *et al.*, 1999). The two metal ions (Mn1 and Mn2) that are proposed to activate the catalytic water molecule (based on comparison to Endo IV) are absent. These results suggest a correlation between the presence of the metal ions and the carboxyl group.

To study whether the carboxyl group can still be present if there are no metal ions, the previously constructed mutant (Paspaleva *et al.*, 2007) E175A was studied in detail. Glu175



**Figure 5.** Structure of UVDE K229L.

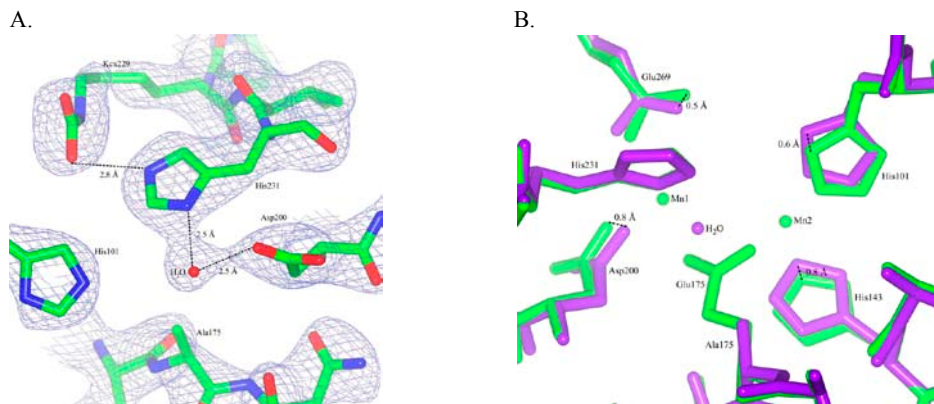
A. Detail of the model and the map of UVDE K229L showing the environment of Leu229. Carbons are depicted in green, oxygens in red and nitrogens in blue. Distances of the water molecule to neighbouring residues are indicated.

B. Detail of a superposition of UVDE K229L (in purple) with the original structure of UVDE (in green). The slight shifts of residues Arg58, Ser99 and His231 are indicated.

bridges Mn1 and Mn2. The mutant E175A was previously found to have no observable activity on DNA containing CPD and a severely reduced activity on DNA containing 6-4PP (5 % activity left). In the structure of this mutant, Mn1 and Mn2 were not observed, as expected. Instead of these ions, there is one water molecule (Figure 6). The metal ion near His244 and His203, Mn3, is present. Most importantly, the modification on Lys229 was still observed. This result shows that the modification can be present in the absence of the Mn1 and Mn2.

Together, the above results suggest that the carboxyl group can be present on Lys229 in absence of Mn1 (as in E175A), but that perhaps this metal ion cannot bind stably in the absence of the carboxyl group (as in structure K229L).

In K229R, Arg229 has moved compared to the carboxylated lysine (Figure 7A and 7B). A water molecule was found at the place where the tip of the carboxylated lysine originally was. Significant shifts are seen for the side-chains of His231 (2.5 Å) and Glu269 (3.2 Å), which have moved outwards away from Arg229. Indeed, His231 moves away from the positive charge of Arg229 and Glu269 moves accordingly. It should be noted that UVDE K229R is actually a double mutant (K229R/L133I) due to a misincorporation during the PCR reaction. No influence of the second mutation on the structure of the protein could be seen (the shift of the C $\alpha$  of residue 133 is only 0.34 Å). Furthermore, Leu133 is a non-conserved residue far away from the DNA binding site, leucine is very similar to isoleucine and, in fact, some bacterial UVDE have isoleucine in this position. Thus, the mutation L133I is likely to have no impact on the behaviour of K229R.



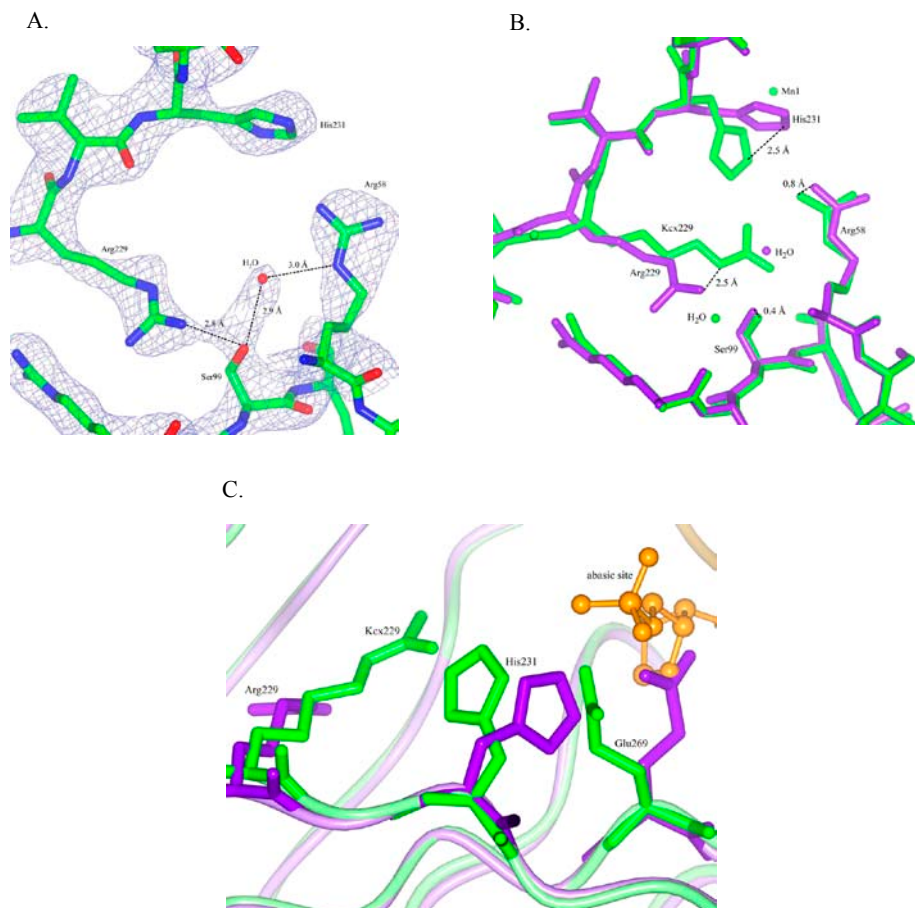
**Figure 6.** Structure of UVDE E175A.

A. Detail of the map (contoured at  $1.25 \sigma$ ) and the model of UVDE E175A showing the environment of Ala175 and the presence of a carboxylated lysine. Carbons are depicted in green, oxygens in red and nitrogens in blue. Distance between an oxygen atom of Kcx229 and His231 is indicated as well as the distances to the water taking the place of Mn1 and Mn2 to neighbouring residues.

B. Detail of the superposition of the original structure of UVDE (in green) and UVDE E175A (in purple) showing the surroundings of residue Glu175/Ala175.

Metal ions were not observed in the K229R structure, which is probably caused by the low pH of the crystallisation condition (below the  $pK_a$  of histidine side-chains). Although two of the metal-coordinating residues have shifted considerably compared to the wild type structure (His231, Glu269), the coordination environments for the three metal ions seem to be intact. Thus, the metals probably can bind stably in all three sites (as the activity also suggests) if the pH permits it, although Mn1 would have to shift over about  $1 \text{ \AA}$  to get into a proper coordination environment.

To get insight into the possible influence of the shifts of His231 and Glu269 on the activity of UVDE K229R, DNA was modelled into K229R based on the crystal structure of endonuclease IV with DNA (Hosfield *et al.*, 1999). In this model, it can be seen that the observed shifts are near the part of the DNA where the damage is expected to be located. In contrast to wild type, in K229R the shifted Glu269 clashes with the abasic site (damaged site) of this model (Figure 7C), suggesting that the part of DNA near the damage might have to bind slightly different in UVDE K229R compared to wild type.



**Figure 7.** Structure of UVDE K229R and model with DNA.

A. Model and map (contoured at 1.5  $\sigma$ ) of UVDE K229R. Carbons are depicted in green, oxygens in red and nitrogens in blue. Distances between a water molecule and Arg58 and Ser99 are indicated.

B. Detail of the superposition of the original structure of UVDE (in green) and UVDE K229R (in purple) showing the position of Arg229 compared to the carboxylated lysine and the large shift in the position of His231. Also the shifts in the position of Arg58 and Ser99 are indicated.

C. Detail of superposition of the original structure of UVDE (in green) and the structure of UVDE K229R (in purple) in which DNA with an abasic site was modelled (orange) based on a comparison to the structure of endonuclease IV with damaged DNA. Arg229/Kcx229, His231 and Glu269 are depicted in cylinder representation, the abasic site of the DNA is depicted in ball-and-stick representation while the rest of the protein and the DNA is depicted in ribbon representation.

## DISCUSSION

The results from mass-spectrometry and crystallisation at low pH showed that the modification we initially observed on Lys229 is (acid) labile. Together with the observed hydrogen bonding pattern of the modification in the original structure, we conclude that the modification is a carboxyl group. Carboxylated lysines have been observed before in prokaryotic proteins and carboxylation of a lysine does not require an enzyme. Therefore, it is a likely possibility for a protein (Tth-UVDE) overexpressed in a foreign host (*E. coli*). With activity assays and structural analysis on mutants, we conclude the carboxyl group might be involved in the stable binding of Mn1 by donating some of its negative charge to the His231. Reduced stability of metal binding might explain the reduced activity of K229A and K229L. For K229L, this reduction in activity is stronger on CPD than on 6-4PP, possibly because the presence of DNA with a 6-4PP might facilitate metal-binding and thus allow activity on 6-4PP. Notably, also UVDE E175A (in which also reduced binding of Mn1 is involved) has a more severe phenotype on CPD than on 6-4PP (Paspaleva *et al.*, 2007).

The mutant K229R, however, does not have a negative charge near His231. Quite the contrary: the protein has a positive charge at position 229 (an arginine). Still, the mutant is active on CPD, 6-4PP and abasic sites with a higher incision efficiency than wild type on these substrates. This increased activity of K229R might be explained by considering a potential mechanism for UVDE where, the protein first recognises a general distortion in the DNA and binds to it (proposed previously based on the broad substrate specificity of *S. pombe* UVDE (Avery *et al.*, 1999). K229R performs this step similar to UVDE wild type, since overall DNA binding was seen not to be affected by the mutation. Then, the DNA gets into the right conformation for incision and the incision takes place. This step is performed more efficiently in K229R than wild type. Structural changes were observed in the active site of K229R and the induced changes block the previously proposed damage-binding pocket. However, the structure of UVDE in complex with DNA is needed to clarify the relevance of the latter observation. The structural changes in the active site of K229R may cause the portion of the DNA near the damage to bind slightly different in the active site, perhaps in a conformation more favourable for incision.

The results on mutant UVDE K229R raise an interesting question. Why did *T. thermophilus* not use an arginine at position 229 if this protein is more efficient than wild type? An answer to this may be found in the activity of UVDE K229R on undamaged DNA. Both UVDE wild type and K229R were found to incise undamaged DNA, which has not been reported for UVDE from *S. pombe*. Apparently, Tth-UVDE has more difficulties in discriminating damaged and undamaged DNA. The high temperature used for incision assays with this thermophile, might play a role in the incision of undamaged DNA, since the DNA can be more readily distorted at higher temperature. Interestingly, we found that the activity on undamaged DNA was much

higher for K229R than for UVDE wild type. Perhaps Arg229 is not favourable for *T. thermophilus* because the incision on undamaged DNA in the cell might be too high.

UVDE is present in several organisms, both prokaryotic and eukaryotic. Is carboxylation a general phenomenon in all these proteins? Two of the residues near the carboxylated lysine, Arg58 and His231, are fully conserved in all UVDE proteins (Golemi *et al.*, 2001). In eukaryotic UVDEs, Lys229 itself is also conserved and these proteins have a threonine instead of a serine at the position corresponding to Ser99. Thus, these proteins might have a similar site at this position and therefore, they might be carboxylated like Tth-UVDE.

In the prokaryotic UVDEs, however, Lys229 is only partially conserved. Most UVDE proteins have a lysine at this position as well as a serine at the position corresponding to Ser99 and thus might be carboxylated like Tth-UVDE. A few UVDEs, however, have a leucine, isoleucine, methionine, glutamic acid, threonine or a valine at the position corresponding to Lys229. Thus, these UVDE homologues cannot have the same modification.

We have shown here that a leucine at position 229 results in an inactive enzyme for Tth-UVDE while some of the other UVDEs do have a leucine at that position, such as *Desulfotalea psychrophila* UVDE (Goosen and Moolenaar, 2008). Perhaps in these UVDEs the environment near His231 is such that a negative charge is provided to His231 by other residues nearby so that stable metal-binding is assured in these proteins as well. Indeed, the UVDEs with a leucine at position 229 all have a glutamic acid or an aspartic acid at the position of Met267, which is spatially close to His231 (the side chain of glutamic acid is 2.5 Å away from the side-chain of His231 if glutamic acid is modeled at this position in the structure of Tth-UVDE). Such a negatively charged residue is not present in the UVDEs that do have a lysine at position 229, those all have a methionine at that position. Moreover, most of the other UVDEs that do not have a lysine at the position of Lys229, also have a glutamic acid near or at position 229. Thus, also these UVDEs might have a negative charge near His231 for assuring proper metal binding, though no definite conclusions can be made in absence of structural data on these proteins.

In conclusion, we think that UVDE from *T. thermophilus* is carboxylated at Lys229. The carboxyl group might be required for stable metal-binding and perhaps also for preventing an unfavourably high incision of undamaged DNA.

**REFERENCES**

- Abendroth J, Niefind K, and Schomburg D, X-ray structure of a dihydropyrimidinase from *Thermus sp.* at 1.3 Angstrom resolution, *J. Mol. Biol.* **320** (2002), pp. 143-156
- Avery AM, Kaur B, Taylor J, Mello JA, Essigmann JM, and Doetsch PW, Substrate specificity of ultraviolet DNA endonuclease (UVDE/UveIp) from *Schizosaccharomyces pombe*, *Nucleic Acids Res.* **27** (1999), pp. 2256-2264
- Bowman KK, Sidik K, Smith CA, Taylor JS, Doetsch PW and Freyer GA, A new ATP-independent DNA endonuclease from *Schizosaccharomyces pombe* that recognizes cyclobutane pyrimidine dimers and 6-4-photoproducts, *Nucleic Acids Res.* **22** (1994), pp. 3026-3032
- Cha J and Mobashery S, Lysine N-zeta-decarboxylation in the BlaR1 protein from *Staphylococcus aureus* at the root of its function as an antibiotic sensor, *J. Am. Chem. Soc.* **129** (2007), pp. 3834-3835
- Collaborative Computational Project 4, The CCP4 suite: Programs for protein crystallography, *Acta Cryst. D* **50** (1994), pp. 760-763
- Diederichs K and Karplus PA, Improved R-factors for diffraction data analysis in macromolecular crystallography, *Nat. Struct. Biol.* **4** (1997), pp. 269-275
- Emsley P and Cowtan K, COOT: Model-building tools for molecular graphics, *Acta Cryst. D* **60** (2004), pp. 2126-2132
- Golema-Kotra D, Cha JY, Meroueh SO, Vakulenko SB and Mobashery S, Resistance of beta-lactam antibiotics and its mediation by the sensor domain of the transmembrane BlaR signalling pathway in *Staphylococcus aureus*, *J. Biol. Chem.* **278** (2003), pp. 18419-18425
- Golemi D, Maveyraud L, Vakulenko S, Samama J and Mobashery S, Critical involvement of a carbamylated lysine in catalytic function of class D beta-lactamases, *Proc. Natl. Acad. Sci. USA.* **98** (2001), pp. 14280-14285
- Goosen N and Moolenaar GF, Repair of UV damage in bacteria, *DNA Repair (Amst.)* **7** (2008), pp. 353-379
- Hosfield DJ, Guan Y, Haas BJ, Cunningham RP and Tainer JA, Structure of the DNA repair enzyme endonuclease IV and its DNA complex: double-nucleotide flipping at basic sites and three-metal-ion catalysis, *Cell* **98** (1999), pp. 397-408
- Iwai S, Chemical synthesis of oligonucleotides containing damaged bases for biological studies, *Nucleotides Nucleic Acids* **25** (2006), pp. 561-582
- Laskowski RA, MacAuthor MW, Moss DS and Thornton M, PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallog.* **261** (1993), pp. 283-291
- Leslie AG, Integration of macromolecular diffraction data, *Acta Cryst. D* **55** (1999), pp. 1696-1702
- Li J, Cross JB, Vreven T, Meroueh SO and Schlegel HB, Lysine carboxylation in proteins: OXA-10 beta-lactamase, *Proteins* **61** (2005), pp. 246-257
- Murshudov GN, Vagin AA and Dodson EJ, Refinement of macromolecular structures by the maximum-likelihood method, *Acta Cryst. D* **50** (1997), pp. 760-763
- Paspaleva K, Thomassen E, Pannu NS, Iwai S, Moolenaar GF, Goosen N and Abrahams JP, Crystal structure of the DNA repair enzyme UV damage endonuclease, *Structure* **15** (2007), pp. 1316-1324
- Perrakis A, Morris R and Lamzin VS, Automated protein model building combined with iterative structure refinement, *Nat. Struct. Biol.* **6** (1999), pp. 458-463
- Potterton L, McNicholas S, Krissinel E, Gruber J, Cowtan K, Emsley P, Murshudov GN, Cohen S, Perrakis A and Noble M, Developments in the CCP4 molecular graphics project, *Acta Cryst. D* **60** (2004), pp. 2288-2294

- Popov AN and Bourenkov GP, Choice of data-collection parameters based on statistical modelling, *Acta Cryst. D* **59** (2003), pp. 1145-1153
- Sidik K, Lieberman HB and Freyer GA, Repair of DNA damaged by UV-light and ionizing-radiation by cell-free-extracts prepared from *Schizosaccharomyces pombe*, *Proc. Natl. Acad. Sci.* **89** (1992), pp. 1212-1216
- Thoden JB, Philips GN, Neal TM, Raushel FM and Holden HM, Molecular structure of dihydroorotase: A paradigm for catalysis through the use of a binuclear metal center, *Biochem.* **40** (2001), pp. 6989-6997
- Thoden JB, Marti-Arbona R, Raushel FM, Holden HM, High-resolution X-ray structure of isoapartyl dipeptidase from *Escherichia coli*, *Biochem.* **42** (2003), pp. 4874-4882
- Vagin A and Teplyakov A, MOLREP: an automated program for molecular replacement, *J. Appl. Cryst.* **30** (1997), pp. 1022-1025
- Verhoeven EE, Van Kesteren M, Turner JJ, Van der Maarel GA, Van Boom JH, Moolenaar GF and Goosen N, The C-terminal region of *Escherichia coli* UvrC contributes to the flexibility of the UvrABC nucleotide excision repair system. *Nucleic Acids Res.* **30** (2002), pp. 2492-2500
- Weiss MS and Hilgenfeld R, On the use of the merging R factor as a quality indicator for X-ray data, *J. Appl. Cryst.* **30** (1997), pp. 203-205



# Damage recognition by UV damage endonuclease from *Schizosaccharomyces pombe*

Chapter

4

Keti Paspaleva, Geri F. Moolenaar and Nora Goosen

---

Laboratory of Molecular Genetics, Leiden Institute of Chemistry, Gorlaeus Laboratories,  
Leiden University, Einsteinweg 55, 2300 RA Leiden, The Netherlands

## SUMMARY

UV damage endonuclease (UVDE) from *Schizosaccharomyces pombe* initiates repair of UV lesions and abasic sites by nicking the DNA 5' to the damaged site. In this paper we show that in addition UVDE incises DNA containing a single-strand nick or gap, but that the enzymatic activity on these substrates as well as on abasic sites strongly depends on the presence of a neighbouring pyrimidine residue. This indicates that, although UVDE may have been derived from an ancestral AP endonuclease its major substrate is a UV lesion and not an AP site. We propose that UVDE rotates two nucleotides into a pocket of the protein in order to bring the scissile bond close to the active site and that purine bases are excluded from this pocket. We also show that in the DNA complex residue Tyr358 of UVDE penetrates the DNA helix causing unstacking of two residues opposite the lesion, thereby stabilizing the protein-DNA interaction, most likely by promoting bending of the DNA. In the absence of Tyr358 the enzyme exhibits an increased catalytic activity on UV-induced lesions, but only at a lower pH of 6.5. At physiological conditions (pH 7.5) the mutant protein completely loses its catalytic activity although it can still bind to the DNA. We propose that in addition to stabilizing the bend in the DNA the hydrophobic side chain of Tyr358 shields the active site from exposure to the solvent.

## INTRODUCTION

UV damage endonuclease (UVDE) is a repair protein, which has been reported to remove a variety of structurally different DNA lesions, including UV photoproducts (CPD and (6-4) PP, apurinic/apyrimidinic (AP) sites and nucleotide mismatches (Kanno *et al.*, 1999, Avery *et al.*, 1999; Kaur *et al.*, 1999). The enzyme nicks the DNA 5' to the damaged nucleotide with the efficiency of the incision highly dependent on the type of lesion. Originally this nuclease was discovered in the fission yeast *Schizosaccharomyces pombe* (Bowman *et al.*, 1994) and in *Neurospora crassa* (Yajima *et al.*, 1995), but later a number of homologues were identified in other fungi, and several eubacterial and archaeobacterial species (Takao *et al.*, 1996, Earl *et al.*, 2002 and Goosen and Moolenaar, 2008).

Recently the structure of UVDE from *T. thermophilus* has been solved (Paspaleva *et al.*, 2007), revealing a high structural similarity with the DNA repair enzyme Endonuclease IV (Endo IV), which participates in *Escherichia coli* Base Excision Repair (BER) by nicking the DNA at the 5' side of an AP site (Hosfield *et al.*, 1999). Both proteins have a TIM-barrel fold and contain three metal ions in their active sites. For Endo IV, however, these metals were shown to be Zn<sup>2+</sup> ions, whereas UVDE uses Mn<sup>2+</sup> ions for catalysis (Paspaleva, manuscript in preparation). Since Endo IV has been crystallised with a substrate containing an AP site, significant light was shed on the mechanism by which the enzyme specifically recognises AP sites. The DNA was seen to be bent by approximately 90°, and both the AP site and the opposing nucleotide were flipped out of the DNA duplex. The flipped-out abasic nucleotide is anchored into a pocket of the enzyme where the 5' phosphate is donated to the three Zn<sup>2+</sup> ions. A normal nucleotide would sterically be excluded from this active site pocket, explaining the AP site selectivity. The opposite base is flipped away from the enzyme and completely solvent exposed. This flipped-out nucleotide is mainly stabilized by the enzyme-induced compression of the DNA backbone and not by direct interaction with the flipped-out base.

The structure of UVDE shows a deep groove that can accommodate kinked DNA similar to Endo IV. The three metal ions at the bottom of this groove are also at similar positions as in Endo IV (Paspaleva *et al.*, 2007), implying that also in UVDE the damaged site must be flipped into a protein pocket in order to bring the scissile bond close to the metal ions. In case of the UVDE enzyme however, this pocket should be able to accommodate CPD's, (6-4)PP's as well as an abasic site.

In this paper we take a closer look at damage recognition by *S. pombe* UVDE by comparing the binding and incision efficiencies on different DNA substrates. Since the full-length protein is unstable (Takao *et al.*, 1996) we used a 228 residue N-terminal truncation of UVDE, which has already been described to be highly active (Takao *et al.*, 1996). We show that besides abasic sites, UVDE also recognises single-strand nicks and gaps, but for all three types of lesion the efficiency is highly dependent on the presence of adjacent pyrimidines. In addition by using a

fluorescent adenine analogue 2-aminopurine (2-AP) we give evidence that binding of UVDE to its DNA substrate causes significant destacking of the bases opposite the lesion.

## MATERIALS AND METHODS

### Proteins

The UVDE gene fragment encoding residues 229 - 599 was amplified by PCR and cloned into the *Bam*HI and *Bst*EII restriction sites of the expression vector pET16b (Novagen), generating an N-terminal fusion of the truncated gene to a 10 X His tag. The resulting plasmid (pETUVDE $\Delta$ 228) was introduced into *E. coli* BL21/codon+ cells (Studier *et al.*, 1990) and the  $\Delta$ 228-UVDE protein purified from a 2 l culture 3 h after induction by IPTG. The cells were lysed by sonication in 6 ml lysis buffer (50 mM Tris pH 7.5, 150 mM NaCl, 10 mM  $\beta$ -mercaptoethanol, 10 % glycerol, 1 % Triton X-100) and separated into a soluble and an insoluble fraction by centrifugation at 37,000 rpm for 30 minutes. The supernatant was loaded on a HiTrap-chelating Ni column, equilibrated with 20 mM Tris pH 7.5 containing 20 mM imidazole and the protein was eluted with a gradient of 20 mM to 250 mM imidazole in the same buffer. Pooled fractions of  $\Delta$ 228-UVDE were loaded on a hydroxyapatite column, which was equilibrated with 10 mM  $KPO_4$  (pH 6.5). Subsequently, the protein was eluted using a gradient from 200 mM to 400 mM  $KPO_4$  (pH 6.5). For further purification, the UVDE fractions were applied to a P11 phosphocellulose column, equilibrated with 300 mM  $KPO_4$  (pH 6.5) and eluted with 1 M NaCl. Finally, the UVDE containing fractions were applied to a Nap5 gel filtration column and eluted with 20 mM Tris pH 7.5, 150 mM NaCl, 10 % glycerol. The UVDE containing fractions showed more than 95 % purity.

The Y358A point mutation was constructed by PCR and verified by sequencing. The mutant UVDE protein was purified using the same purification procedure as described for the wild type enzyme and showed the same elution/purification profile.

The Endonuclease IV (Endo IV) and Endonuclease III (Endo III) enzymes were obtained commercially (New England Biolabs).

### DNA substrates

The 30 bp DNA substrates used in this study are summarised in Table 1. The oligonucleotides containing CPD or (6-4)PP lesions were synthesised as described (Iwai, 2006). The 30 bp substrates containing an abasic site (via incorporation of a tetrahydrofuran-dSpacer), thymine glycol (TG) or 2-AP were obtained commercially (Eurogentec, Belgium). The DNA substrates were 5' radioactively labelled using polynucleotide kinase as described (Verhoeven *et al.*, 2002). For 3' labelling the damaged top strands were annealed to a corresponding bottom strand with one additional G residue at its 5' end. Subsequently the DNA fragments were incubated with Klenow polymerase and  $\alpha$ -<sup>32</sup>P-dCTP as described (Moolenaar *et al.*, 2005).

**Incision assay**

The labelled DNA substrates were incubated with the indicated amount of UVDE in a 20  $\mu$ l reaction mix (20 mM HEPES pH 6.5, 100 mM NaCl, 10 mM MgCl<sub>2</sub>, 1 mM MnCl<sub>2</sub>). After 15 minutes incubation at 30°C the reaction was terminated by adding 3  $\mu$ l of 0.33 M EDTA, 3.3 % SDS and 2.4  $\mu$ l glycogen (4  $\mu$ g/ $\mu$ l), followed by an ethanol precipitation. For measuring the incision percentages the samples were run on a small (8 cm) 15 % acrylamide gel. For determination of the incision positions longer gels were used (20 cm) allowing separation of the different incision products.

Incision with Endo IV was performed in a 20  $\mu$ l reaction mix in the presence of 50 mM Tris-HCl pH 7.9, 100 mM NaCl, 10 mM MgCl<sub>2</sub>, 1 mM DTT) for 15 minutes at 37 °C. Incision with Endo III was performed for 15 minutes at 37°C in buffer supplied by the manufacturer.

**Filter binding Assay**

The filter binding assays were conducted in 20  $\mu$ l samples containing 5 nM of UVDE and 4 nM of the terminally labelled DNA substrates in a reaction buffer containing 20 mM Tris pH 6.5, 100 mM NaCl. The mixes were incubated for 10 minutes at 30°C. If needed 1 mM MnCl<sub>2</sub> and 10 mM MgCl<sub>2</sub> were included in the reaction buffer. After the incubation 0.5 ml of preheated (30°C) reaction buffer was added to each sample. The mixture was then poured over a nitrocellulose filter (Millipore 0.45  $\mu$ m HA) and the filter was rinsed three times with 0.5 ml of the reaction buffer. Each sample was corrected for the amount of DNA retained on a filter in the absence of the protein. Binding is expressed as the percentage of input DNA retained on the filter by the enzymes. To determine the binding properties of UVDE under the conditions used for the 2-aminopurine measurements 4 nM of the 5' terminally labelled substrates were mixed with 0.5  $\mu$ M unlabelled DNA and incubated with 2.5  $\mu$ M UVDE for 10 min. at 30°C in a total volume of 30  $\mu$ l.

**2-AP measurements**

60  $\mu$ l samples containing 0.5  $\mu$ M DNA and 2.5  $\mu$ M UVDE were incubated in 20 mM Tris buffer (pH 6.5) for 10 min at 30°C. If needed 1 mM MnCl<sub>2</sub> and 10 mM MgCl<sub>2</sub> were included in the incubation mixture. Where indicated 300 mM acrylamide was added to the samples after incubation. The samples were transferred to a 3 mm x 3 mm quartz cuvette and placed in the fluorimeter. Fluorescence emission spectra were obtained using a PerkinElmer LS 50B fluorimeter, connected to a temperature variable water bath to maintain a temperature inside the cuvette of 30°C. The excitation wavelength was set at 310 nm, and the emission spectra were obtained by scanning from 325 to 475 nm. The excitation and emission slit widths were 5 nm and 10 nm, respectively. All spectra presented are the average of at least to experiments and were corrected for the spectra of equivalent incubation mixtures without DNA. The free 2-aminopurine signal was obtained by measuring 0.5  $\mu$ M 2-aminopurine riboside-3',5'-cyclic monophosphate (BioLog) in 60  $\mu$ l of 20 mM Tris buffer at 30°C.

## UV survival test

The *Xba*I - *Bam*HI fragment from pETUVDEA228 (with or without the Y358A mutation) was inserted in pIC-19R, resulting in expression of the (mutant) UVDE under control of the Plac promoter. The plasmids were introduced in CS 5018 ( $\Delta$ *uvrA*,  $\Delta$ *uvrB*) (Moolenaar *et al.*, 2005). Cells were grown in LB supplemented with 40  $\mu$ g/ml ampicillin and 1 mM IPTG until OD<sub>600</sub> reached 0.3. Subsequently 1  $\mu$ l drops of a 10<sup>-1</sup> dilution were spotted on an LB plate containing 40  $\mu$ g/ml ampicillin and 1 mM IPTG. The drops were irradiated with the indicated dose of UV and plates were incubated overnight at 30°C.

## RESULTS

### Incision of an abasic site is dependent of the flanking bases

The structural similarities between UVDE and Endo IV suggest that both proteins bind the damaged DNA in a similar fashion. They both share a wide groove housing the active site

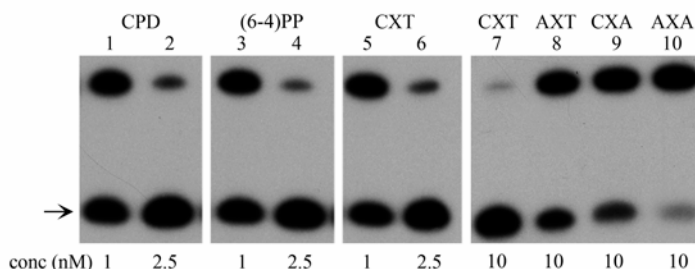
**Table 1.** DNA substrates used in this study.

no damage	5' CTCGTCAGCATCTTCATCATAACAGTCAGTG 3'
	3' GAGCAGTCGTAGAAGTAGTATGTCAGTCAC 5'
CPD and (6-4)PP	5' CTCGTCAGCAT <b>TT</b> TCATCATAACAGTCAGTG 3'
	3' GAGCAGTCGTAGAAGTAGTATGTCAGTCAC 5'
CXT abasic site	5' CTCGTCAGCAT <b>CX</b> TCATCATAACAGTCAGTG 3'
	3' GAGCAGTCGTAGAAGTAGTATGTCAGTCAC 5'
AXT abasic site	5' CTCGTCAGCATA <b>AX</b> TCATCATAACAGTCAGTG 3'
	3' GAGCAGTCGTATAAGTAGTATGTCAGTCAC 5'
CXA abasic site	5' CTCGTCAGCAT <b>CX</b> ACATCATAACAGTCAGTG 3'
	3' GAGCAGTCGTAGATGTAGTATGTCAGTCAC 5'
AXA abasic site	5' CTCGTCAGCATA <b>AX</b> ACATCATAACAGTCAGTG 3'
	3' GAGCAGTCGTATATGTAGTATGTCAGTCAC 5'
C*T ss-nick	5' CTCGTCAGCAT <b>CT</b> TCATCATAACAGTCAGTG 3'
	3' GAGCAGTCGTAGAAGTAGTATGTCAGTCAC 5'
C*A ss-nick	5' CTCGTCAGCAT <b>CA</b> TCATCATAACAGTCAGTG 3'
	3' GAGCAGTCGTAGTAGTATGTCAGTCAC 5'
A*T ss-nick	5' CTCGTCAGCATA <b>AT</b> TCATCATAACAGTCAGTG 3'
	3' GAGCAGTCGTATAAGTAGTATGTCAGTCAC 5'
A*A ss-nick	5' CTCGTCAGCATA <b>AA</b> TCATCATAACAGTCAGTG 3'
	3' GAGCAGTCGTATTAGTAGTATGTCAGTCAC 5'
T*T ss-nick	5' CGTGTGAGGTCGTTCTGAGGTT <b>TT</b> TTTGTAAATGTGCCGTAAGTAATCCC 3'
	3' GCACACTCCAGCAAGACTCCAAAAAACATTACACGGGCATTCATTAGGG 5'
1 nt gap	5' AGTTCTATGCGCACCGAATTCCTCACT GAACCCAAGCTTGCCGGGCTCT 3'
	3' TCAAGATACGCGTGGCTTAAGGGTGACCTTGGGTTTCAACGGCCCGGAGA 5'
Thymine glycol	5' CTCGTCAGCAT <b>T</b> TCATCATAACAGTCAGTG3'
	3' GAGCAGTCGTAGAAGTAGTATGTCAGTCAC5'

The positions of the CPD (**TT**), (6-4)PP (**TT**), AP site (**X**) and thymine glycol (**T**) are indicated in bold. The positions of the single-stranded nicks in fragments C\*T, A\*A and T\*T are between the two bold residues.

deeply at its bottom and extensive positive charges positioned at both ends of the groove. Endo IV has been shown to flip its target, the abasic site into an extrahelical position inside a pocket of the protein. UVDE however, recognizes CPD dimers and (6-4)PPs where both thymines are covalently linked. This implies that to be able to bring the scissile phosphodiester bond close to the metals of its active site it must flip the two crosslinked nucleotides into a protein pocket. This poses the question whether upon binding of an abasic (AP) site only this AP site will be flipped into the proposed protein pocket or the AP site in consortium with its neighbouring nucleotide, similar to the crosslinked pyrimidines. If this latter would be true one might expect that the efficiency of abasic site incision might be influenced by the nature of this flanking base. To test this we used four 30 bp DNA fragments which only differ in the bases flanking the AP site on the 5' or 3' side (substrates CXT, AXT, CXA and AXA, with X representing the abasic site (see Table 1).

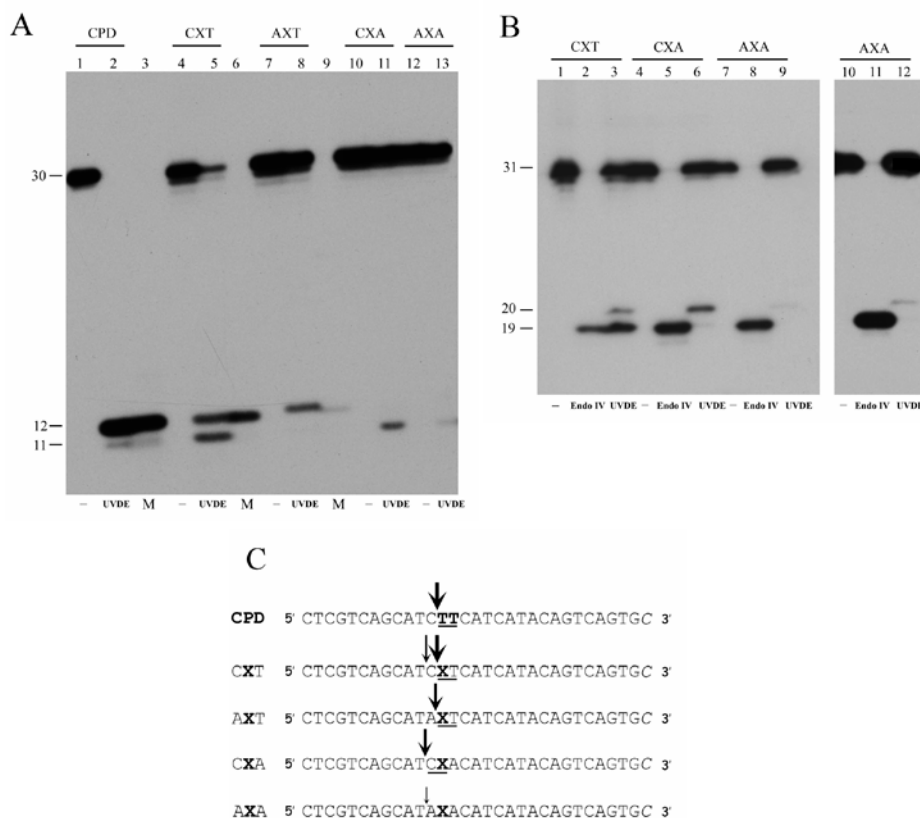
The highest incision was obtained with substrate CXT (Figure 1, lane 7), which has the AP site flanked by two pyrimidines (> 95 %). The incision efficiency on this AP site is similar to the incision on a CPD or (6-4)PP (lanes 1 - 6). Incision is significantly lower on an abasic site in the sequences AXT (20 %) and CXA (15 %) (lanes 8, 9) and extremely low on substrate AXA (< 5 %, lane 10) which has the lesion flanked by two purines. The AP-endonuclease IV enzyme, however, incised all four DNA constructs with equal efficiencies (not shown). Apparently, the incision activity of UVDE on an abasic site indeed strongly depends on the sequence surrounding the lesion.



**Figure 1.** UVDE incision of CPD, (6-4)PP and AP sites in different sequence contexts.

The 5'-terminally labelled substrates (0.1 nM) were incubated with the indicated amount of UVDE for 15 minutes. The different substrates are shown above each lane. The arrow indicates the incision product.

We also determined the incision positions on the different substrates using constructs labelled on either the 5' (Figure 2A) or the 3' side (Figure 2B) of the damaged strand. As a control the incision product of UVDE on a CPD-containing fragment (Figure 2A, lane 2) was included in the gel, showing that on this substrate UVDE uniquely nicks the DNA immediately 5' to the CPD lesion. The minor band that migrates one position lower in the gel can be ascribed to the presence of a small amount of 29-mer in the CPD-substrate (lane 1). With substrate AXT UVDE also incises the DNA immediately 5' to the AP site (Figure 2A, lane 8), but on substrates



**Figure 2.** Determination of incision positions.

A The 5'-terminally labelled substrates (0.8 nM) were incubated without protein (lanes 1, 4, 7, 10, 12) or with 10 nM (lane 2) or 5 nM of UVDE (lanes 5, 8, 11, 13) for 15 minutes and loaded on a long acrylamide gel. The positions of the non-incised fragment (30 nt) and the incision products (11 and 12 nt) are indicated. The substrates used are indicated above the lanes. The lanes containing the 12 nt marker fragment (with the sequence corresponding to the first 12 nt of the top strand of the CPD substrate) are indicated with M.

B The 3'-terminally labelled substrates were incubated without protein (lanes 1, 4, 7, 10), with Endo IV (lanes 2, 5, 8, 11) and 2 nM (lane 3) or 5 nM UVDE (lanes 6, 9, 12) for 15 minutes. Lanes 10 - 12 represent a longer exposure of lanes 7 - 9 to visualize the incision product. The positions of the non-incised fragment (31 nt) and the incision products (19 and 20 nt) are indicated.

C Schematic representation of the incision positions. The arrows indicate the positions of the nicks made by UVDE. The sizes of the arrows correspond to the efficiencies of the direct incisions. The residues that take up the same position with respect to the nick as the CPD are underlined. Note that only the 3'-labelled fragments contain an additional C (italics) at the 3' end.

CXA and AXA the incision position is shifted one nucleotide in the 5' direction (Figure 2A, lanes 11, 13; Figure 2B, lanes 6, 12). With the CXT sequence clearly two incision positions are observed which have equal intensities with the 5'-labelled DNA (Figure 2A, lane 5). The 3' labelling, however, reveals that the major incision is at position 19, which is immediately 5' to the abasic site (Figure 2B, lane 3). There is only a minor incision on the position which is 1 nt shifted to the 5' side, indicating that the 11 nt fragment observed with the 5' labelling is mainly the result of a second incision that is made after introduction of the nick adjacent to the AP site (see also below). When we used Endo IV on the different AP site fragments only one incision product was obtained, as result of incision directly 5' to the AP site (Figure 2B, lanes 2, 5, 8). Apparently for UVDE the bases flanking the AP site not only determine the efficiency of incision but also the position of the nick.

When we align the different (major) incision positions of the AP constructs with that of the CPD substrate (Figure 2C) we can extrapolate which two residues (the AP site and a flanking nucleotide) will take up the same position as the CPD in the protein-DNA complex formed on the abasic sites. These combinations are incised with different efficiencies: XT > CX >> AX. The combination XA apparently does not lead to a productive complex, since neither in the CXA nor in the AXA fragment a significant incision is observed immediately adjacent to the AP site. Taken together our results indicate that indeed UVDE recognizes the abasic site together with a flanking nucleotide either on the 5'- or on the 3' side with a high preference for a pyrimidine as a neighbour. This strongly suggests that like for the crosslinked photoproducts also on the AP site constructs UVDE rotates two residues into a protein pocket, the abasic site together with a flanking pyrimidine. The enzyme shows a preference for having this pyrimidine on the 3' side of the abasic site since on the CXT fragment the majority of the incision is directly 5' to the AP site.

Next we tested the binding of UVDE to the four DNA fragments using a filter binding assay described in Materials and Methods. The binding was performed in the absence (preventing incision) or presence (allowing incision) of 1mM MnCl<sub>2</sub> and 10 mM MgCl<sub>2</sub> (Table 2). In both conditions the percentage of binding indeed differs for the 4 constructs, with CXT giving the highest amount of complex (comparable to the binding of a CPD lesion) and AXA the lowest. For all four substrates, however, the binding is still significantly higher than binding to undamaged DNA (Table 2). Even for the AXA substrate, which is very poorly incised the binding is still 50 % of that of the CXT fragment. Apparently the binding per se does not always lead to a productive complex. Comparing the binding data of the two incubation conditions reveal that in the presence of the metal ions DNA binding is somewhat increased for the fragments that are efficiently incised by UVDE (CPD, (6-4)PP and CXT) but remains unaltered for the substrates that are only poorly incised (CXA, AXA). This suggests that the presence of a nick stabilizes the protein-DNA complex.

**Table 2.** UVDE binding to the different DNA substrates.

Substrate	% binding no cofactors	% binding Mn <sup>2+</sup> + Mg <sup>2+</sup>
no damage	2 ± 1	n.d.
CPD	55 ± 3	62 ± 2
(6-4)PP	57 ± 1	66 ± 1
abasic (CXT)	53 ± 3	59 ± 1
abasic (AXT)	41 ± 2	43 ± 1
abasic (CXA)	38 ± 2	38 ± 1
abasic (AXA)	29 ± 3	27 ± 1
ss-nick (C*T)	49 ± 2	47 ± 4
ss-nick (C*A)	44 ± 2	45 ± 1
ss-nick (A*T)	32 ± 1	30 ± 1
ss-nick (A*A)	22 ± 1	23 ± 1
TG	58 ± 1	60 ± 2

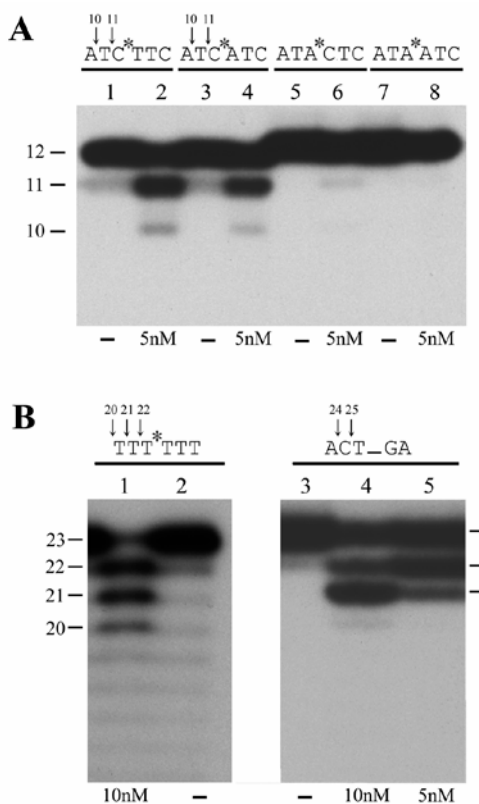
X, position of AP site; \*, position of nick; TG, thymine glycol

Incubation was done with 4 nM DNA and 5 nM UVDE, n.d. not determined.

### UVDE incises DNA containing a nick or a gap

As shown above after nicking of the CXT substrate directly 5' to the abasic site, UVDE makes a second incision 5' to the adjacent C residue. This indicates that formation of a productive complex on the CX sequence is improved by the presence of a nick between the two residues. To test whether for this incision the AP site is still required we also determined the activity of UVDE on the same DNA fragment with a ss-nick but without the abasic site (C\*T). Figure 3A (lane 2) shows that this DNA substrate is efficiently incised by UVDE (approximately 50 %). A DNA fragment with identical sequence without the nick is not incised at all (not shown). The two subsequent incisions on the CXT abasic site fragment also show that under similar incubation conditions approximately 50 % of the first incision product is additionally nicked by UVDE (note the equal intensities of the two incision products in Figure 2A, lane 5). Apparently whether there is a thymine or an abasic site on the 3' side of the nick does not significantly influence the incision. Even with a purine on the 3' side (C\*A, lane 4) a similar amount of incision products is found. DNA fragments with a purine on the 5' side (A\*C and A\*A), however, are hardly incised by UVDE (lanes 6, 8). This means that there is a preference for a pyrimidine at the 5' side of the nick but the nature or even the presence of the base on the 3' side does not seem to be of importance. Concomitant with the incision efficiencies also the DNA binding to the C\*T and C\*A sequences is higher than to the nicks with a 5' flanking purine residue, but again the binding to the A\*C and A\*A fragments is significantly higher than to undamaged DNA (Table 2), as was observed for the abasic site fragment that was very poorly incised (AXA). UVDE incises the C\*T and C\*A fragments 5' of the C that is flanking the nick, but also a second incision product is observed which results from incision 1 nt further downstream (Figure 3A, lanes 2, 4). When using a 50 bp substrate containing a nick flanked by two stretches of pyrimidines (constructed for the fluorescence studies described below) even two additional incision products were obtained

(20 and 21 nt, Figure 3B, lane 1). The intensities of these incision products suggest that they result from subsequent incision events. This would imply that UVDE can also nick DNA 5' to a nucleotide that is flanked by a 1 or 2 nt gap. To test this we constructed a DNA fragment containing a 1 nt gap (Table 1). Indeed UVDE does nick this DNA resulting in two incision products of 25 and 24 nt (Figure 3B, lane 5). At a higher protein concentration the amount of the 25 nt fragment decreases whereas that of the 24 nt fragment increases showing that the 24 nt fragment results from two consecutive incision events. Both incisions are made 5' to a pyrimidine residue. A subsequent third incision is not observed, most likely because this nick would have to be made 5' to a purine residue (see sequence Figure 3B). In the T\*T fragment there are three pyrimidines present explaining why on this substrate three successive incisions



**Figure 3.** Activity of UVDE on a DNA fragment with a ss-nick or 1 nt gap.

A. The 5'-terminally labelled substrates (0.8 nM) were incubated with 5 nM of UVDE for 15 minutes.

B. Similar incubations containing 0.3 nM of T\*T nick fragment (lanes 1, 2) or 0.8 nM of the gapped substrate (lanes 3 - 5) with the indicated amount of UVDE. The relevant sequences of the DNA substrates are shown above the lanes with the position of the nick (\*) and the position of the gap (-) and the incision positions (arrows).

can be made. Taken together the results show that UVDE can also form productive complexes on a pyrimidine residue that lacks a 3' neighbour, implying that on these substrates rotation of only one nucleotide into the proposed protein pocket is sufficient. The fact that on a nicked substrate, in contrast to the abasic site constructs, the nature of the 3' flanking base is of much less importance suggests that because of the interrupted phosphodiester backbone also here only one nucleotide which is 5' of this nick will be rotated. It should be noted that although the binding of the CXT abasic fragment is similar to that of the C\*T nicked DNA, the overall incision of the abasic site is still higher (compare Figure 2A, lane 5 with Figure 3A, lane 2). This implies that the catalytic reaction is more successful in complexes with two flipped bases on an intact DNA backbone than on complexes with one flipped base on a broken DNA backbone.

### UVDE-mediated unstacking of nucleotides opposite the lesion

In the co-crystal structure of Endo IV with DNA not only the AP site but also the opposite base is extruded from the DNA helix (Hosfield *et al.*, 1999). To test whether base flipping in the opposing strand also occurs upon binding of UVDE we made use of the fluorescent adenine analogue 2-AP that we incorporated in the different DNA substrates opposite the lesion (CPD, (6-4)PP, CXT abasic). Excitation of 2-AP with light of a wavelength of 310 nm results in an emission of about 370 nm. In a double-stranded DNA fragment this fluorescence is significantly quenched by base stacking interactions and therefore 2-AP can be used as a spectroscopic probe

**Table 3.** UVDE binding to 2-aminopurine (2-AP) containing DNA substrates.

Damage	2-AP position	% binding wt UVDE	% binding Y358A
No damage	1	5 ± 3	4 ± 2
No damage	2	8 ± 2	6 ± 3
CPD	no	61 ± 1	28 ± 1
CPD	1	60 ± 1	25 ± 1
CPD	2	59 ± 1	27 ± 1
(6-4)PP	no	59 ± 3	60 ± 3
(6-4)PP	1	61 ± 4	62 ± 2
(6-4)PP	2	65 ± 1	66 ± 1
abasic (CXT)	no	62 ± 1	16 ± 1
abasic (CXT)	1	61 ± 4	15 ± 2
abasic (CXT)	2	60 ± 1	15 ± 1
nick (T*T)	no	57 ± 1	n.d.
nick (T*T)	1	57 ± 2	n.d.
nick (T*T)	2	59 ± 1	n.d.
TG	no	60 ± 1	n.d.
TG	1	60 ± 2	n.d.
TG	2	61 ± 1	n.d.

X, position of AP site, TG, thymine glycol. Incubation was done using the same conditions as used for the fluorescence assays without metal cofactors. The positions of the 2-AP residues are as shown in Figures 4 and 5.

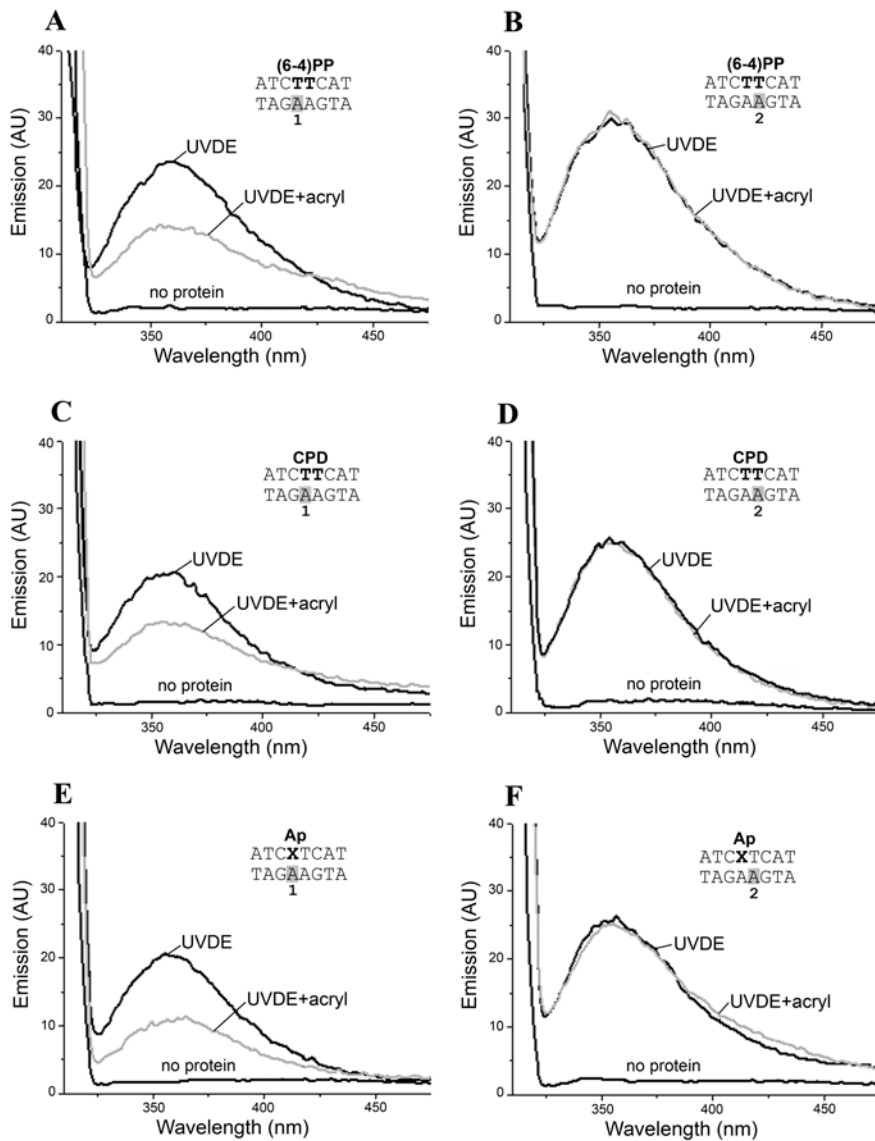
for base unstacking and base flipping (Allan *et al.*, 1996; Christine *et al.*, 2002; Holz *et al.*, 1998; McCullough *et al.*, 1997; Bandwar and Patel, 2001 and Malta *et al.*, 2006). The incision efficiencies of the different substrates were not influenced by the presence of a 2-AP instead of a normal adenine (not shown) and also the DNA binding efficiencies are similar (Table 3).

In the absence of protein none of the DNA fragments exhibit significant fluorescence (Figure 4, A - F), showing that the different lesions do not disturb the base stacking in the opposing strand. The addition of UVDE to a non-damaged DNA fragment containing either of the 2-AP residues also did not give rise to any fluorescence signal (not shown).

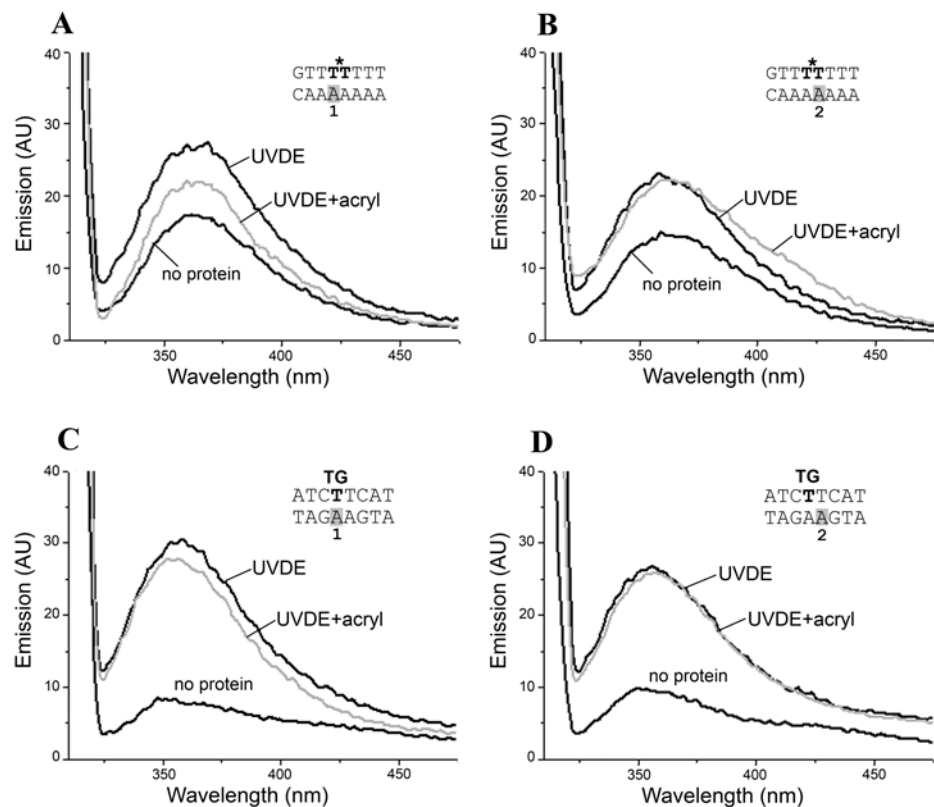
Upon incubation of the (6-4)PP-containing DNA with UVDE in the absence of metal ions (preventing DNA incision) fluorescence of the 2-AP at position 1 (opposite the 5' T) as well as the 2-AP at position 2 (opposite the 3' T) increases significantly to 25 and 32 AU respectively (Figure 4, A and B). Considering that the fluorescence of the same concentration of free 2-AP is around 100 AU and that under conditions used for the fluorescence measurements about 60 % of the DNA is bound (Table 3), significant destacking of the two damage-opposing bases does occur upon UVDE binding. For the CPD lesion similar destacking of the two bases can be detected, again with the signal of the 2-AP opposite the 3' T somewhat higher than its neighbour opposite the 5' T (Figure 4, C and D). Binding of UVDE to the abasic site in the sequence CXT shows that the 2-AP residues opposite the abasic site (position 1) and the 3' flanking T (position 2) show the same fluorescence as the 2-AP's opposite the CPD or (6-4)PP lesions (Figure 4, E and F). This confirms that in this substrate UVDE is positioned on the XT sequence in a similar way as on the CPD and (6-4)PP.

We also tested fluorescence of the T\*T nicked DNA with 2-AP residues opposing the thymines flanking the nick (Figure 5A, B). The T\*T DNA fragment by itself (in the absence of protein) already gives rise to significant fluorescence. Apparently the presence of a nick in the A/T-rich sequence largely enhances the flexibility of this DNA. In the presence of UVDE a similar fluorescence of the 2-AP residues is observed as for the other substrates (Figure 5, A and B) showing that unstacking of the bases in the opposite strand also occurs when the 'damaged' strand is broken.

We also tested whether conformation of the DNA in the complex changes after incision. To this purpose the DNA substrates were pre-incubated with UVDE in the presence of the metal cofactors for 10 minutes to allow incision, after which the fluorescence spectra were recorded (Table 4). Incision of the different substrates under these conditions was confirmed (not shown). For none of the 2-AP residues a significant change in fluorescence could be observed (Table 4), indicating that upon incision the residues opposing the lesion remain in the same unstacked conformation.



**Figure 4.** Fluorescence emission spectra. The DNA fragments (0.5  $\mu$ M) with a (6-4)PP (A and B) a CPD (C and D) or an AP lesion (E and F) were incubated with or without 2.5  $\mu$ M UVDE for 10 minutes at 30°C. After incubation the sample was transferred to a cuvette and emission spectra were recorded at 30°C (excitation at 310 nm). The relevant sequences of the different substrates are shown with the 2-AP residue at position 1 (A, C, E) or position 2 (B, D, F). The spectra of the samples without protein or after incubation with UVDE are in black as indicated. In grey are the spectra recorded in the presence of acrylamide.



**Figure 5.** Fluorescence emission spectra. The DNA fragments (0.5  $\mu$ M) with a ss-nick (T\*T, panels A and B) or thymine glycol (TG, panels C and D) were incubated with or without 2.5  $\mu$ M UVDE for 10 minutes at 30°C. After incubation the sample was transferred to a cuvette and emission spectra were recorded at 30°C (excitation at 310 nm). The relevant sequences of the different substrates are shown with the 2-AP residue at position 1 (A and C) or position 2 (B and D). The spectra of the samples without protein or after incubation with UVDE are in black as indicated. In grey are the spectra recorded in the presence of acrylamide.

**Table 4.** Fluorescence signals before and after incision.

Damage	2-AP position	Fluorescence no incision	Fluorescence 10 min incision
CPD	1	23 $\pm$ 1	25 $\pm$ 4
CPD	2	27 $\pm$ 2	28 $\pm$ 2
(6-4)PP	1	25 $\pm$ 2	27 $\pm$ 3
(6-4)PP	2	32 $\pm$ 1	30 $\pm$ 2

The fluorescence signal (emission at 370 nm) is shown in arbitrary units (AU).

In the Endo IV protein-DNA complex the extrahelical base opposite the AP site does not make any contacts with residues of the protein and is fully solvent exposed. To test whether UVDE extrudes the opposing bases in a similar way we measured fluorescence of the UVDE-DNA

complex in the presence of acrylamide. Acrylamide has been shown to quench the fluorescence of 2-AP but it can only do so if the 2-AP is accessible to the compound, i.e. solvent exposed (Malta *et al.*, 2006, Rai *et al.*, 2003). Figures 4 and 5 (grey lines) show that for all four substrates fluorescence of the 2-AP at position 2 remains unaltered upon addition of acrylamide, indicating that this residue is shielded from solution by residue(s) of the UVDE protein. The 2-AP at position 1 appears not fully protected by the protein as in all four substrates it is partly quenched by the acrylamide (Figures 4 and 5). For the T\*T containing fragment the quenching of 2-AP at position 1 seems less than for the other substrates (Figure 5A), but one should bear in mind that for this particular substrate the unbound DNA that is still present in the sample also contributes to the fluorescence signal. In any case, unlike Endo IV for none of the DNA substrates the residues in the complementary strand become fully solvent exposed upon UVDE binding.

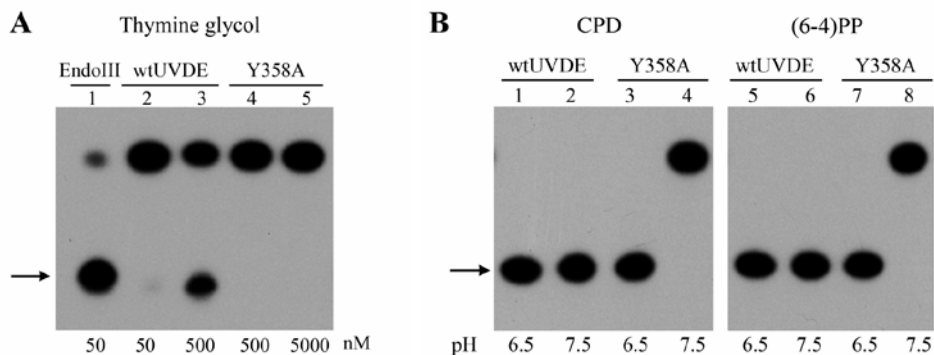
### **UVDE recognizes but does not efficiently incise DNA containing a thymine glycol.**

It has been reported that UVDE does not incise DNA containing a thymine glycol (Kanno *et al.*, 1999), but since the data of this experiment were not included in the paper, this might have been due to the presence of purine residues flanking the thymine glycol (TG) lesion. We therefore constructed a DNA fragment in which the TG is flanked by pyrimidine residues in the same sequence context as the optimal abasic site substrate CXT (Table 1). The binding of UVDE to this DNA fragment is indeed comparable to that of the CXT, CPD or (6-4)PP containing fragments (Table 2). Also under conditions of the 2-AP fluorescence experiments a similar binding is observed (Table 3) and these complexes do give rise to fluorescence of the 2-AP residues in the opposing strand (Figure 5, C and D). The addition of acrylamide again shows that these 2-AP residues do not become solvent exposed. The quenching at position 1 is somewhat less compared to the other types of lesion, indicating that the conformation of this residue might be different, shielding it more from the solution.

When we tested the incision of the TG-containing DNA by UVDE the nicking activity appeared surprisingly low. At the 'normal' concentrations (5 nM) UVDE does hardly incise the DNA and only at a 10 times higher concentration some nicking (20 %) is observed (Figure 6A, lanes 2, 3). As a control the Endo III enzyme does efficiently induce incision on the same substrate (Figure 6A, lane 1). Apparently although UVDE does bind the TG lesion and also induces destacking of the bases in the opposing strand, it cannot effectively position the scissile bond into the active site of the protein.

### **Residue Y358 stabilizes the extra helical bases opposite the lesion.**

In the crystal structure of *T. thermophilus* UVDE an aromatic residue (Tyr105) was seen to point out from the proposed DNA binding groove into the solvent (Paspaleva *et al.*, 2007). The importance of Tyr105 for the activity of the *T. thermophilus* UVDE became clear when this residue was changed into an alanine, since the incision activity of the resulting mutant was severely reduced (Paspaleva *et al.*, 2007).



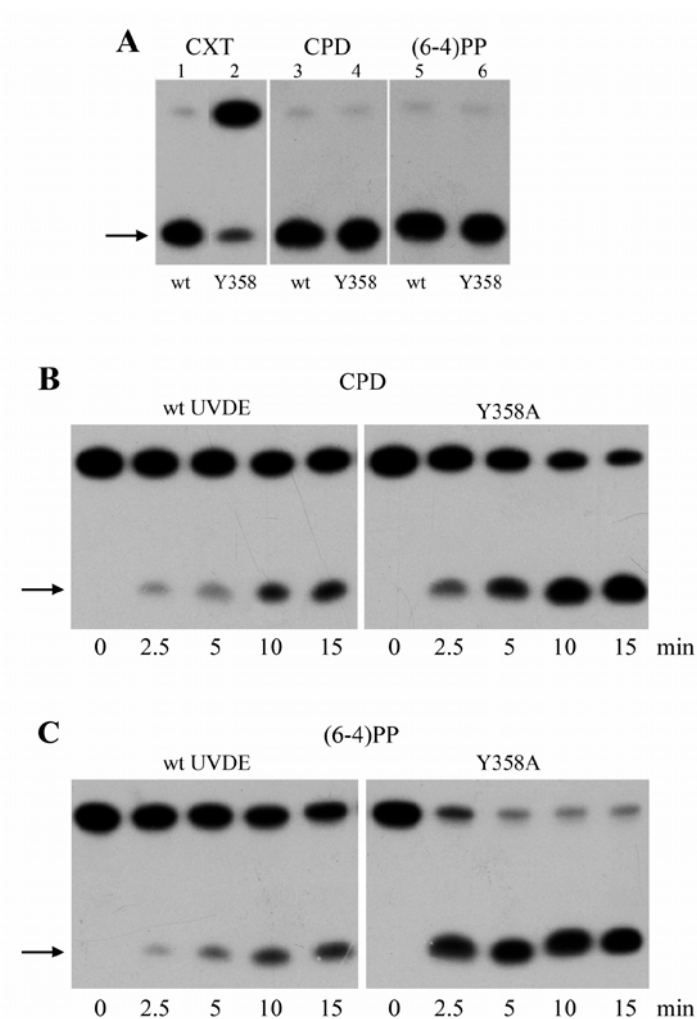
**Figure 6.** Incision activity on thymine glycol and effect of pH on incision.

A. The 5'-labelled 30 bp DNA fragment (0.2 nM) with a TG lesion was incubated with the indicated amount of Endo III (lane 1) and wtUVDE (lanes 2 and 3).

B. The 5'-labelled 30 bp DNA fragment (0.2 nM) with a CPD or (6-4)PP was incubated with 5 nM of UVDE at two different pH conditions as shown. The arrow indicates the position of the incision product.

To our surprise, mutating the equivalent tyrosine residue in *S. pombe* UVDE (Y358A) resulted in a protein that was severely reduced in incision of DNA containing an AP site (Figure 7A, lane 2), but the incision on the CPD or (6-4)PP was still very efficient (Figure 7A, lanes 3 - 6). Filter binding studies revealed that indeed the binding of the mutant protein to an AP site is significantly reduced (Table 3). The binding of Y358A to the (6-4)PP is similar to that of the wild type protein, but also the binding to the CPD appears reduced (Table 3). A kinetic assay with a limiting amount of protein revealed that the incisions of the CPD and (6-4)PP by the mutant protein are even much faster than for the wild type protein (Figure 7, B and C). The (6-4)PP containing DNA is more efficiently nicked by the mutant protein than the CPD, which probably reflects the difference in binding between these substrates. Apparently the absence of Tyr358 does reduce the stability of the protein-DNA complex on the CPD, but this is largely compensated by a much faster incision step. For the (6-4)PP Tyr358 is not required for stable binding and also for this DNA fragment the removal of this residue speeds up the incision reaction.

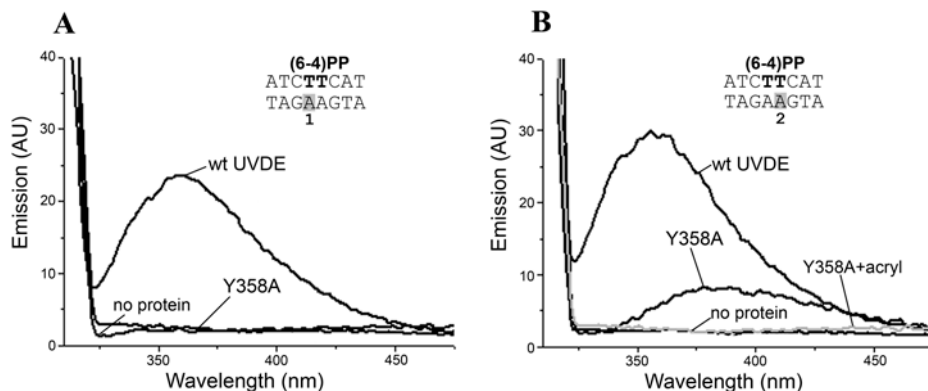
Next we tested the effect of the Y358A mutation on the 2-AP fluorescence. For this we used the DNA with the (6-4)PP lesion, since under the conditions used for the fluorescence assay the mutant protein forms a similar amount of complexes on this lesion as the wild type protein (Table 3). Position 1 no longer shows any fluorescence upon binding of the mutant protein (Figure 8A) and fluorescence at position 2 is reduced by about a factor of 3 (Figure 8B). The peak of the remaining fluorescence signal at position 2 is significantly shifted to a higher wavelength indicating that the 2-AP residue in the mutant complex is in a different environment. This was confirmed by the observation that the remaining fluorescence at position 2 can now



**Figure 7.** Incision activity of the Y358A mutant.

A. The 5'-labelled 30 bp DNA fragments (0.2 nM) with a CPD, (6-4)PP or AP site (CXT) were incubated with 5 nM UVDE or mutant Y358 for 15 minutes at 30°C (B and C). The CPD fragment (0.1 nM) and (6-4)PP fragment (0.1 nM) were incubated with 0.25 nM of (mutant) UVDE for the indicated amount of time (minutes). The incision products are indicated with an arrow.

be fully quenched by acrylamide (Figure 8B), whereas the fluorescence of the same residue in the complex formed by the wild type protein was fully protected from quenching (Figure 4B). This clearly shows that residue Tyr358 not only stabilizes the destacking of the bases in the non-damaged strand but also that in the absence of Tyr358 position 2 becomes more solvent exposed.



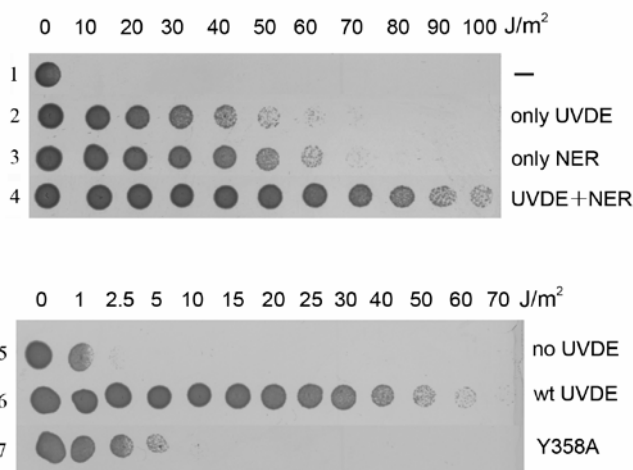
**Figure 8.** Fluorescence emission spectra of wt UVDE and mutant Y358A on a substrate containing a (6-4)PP.

The position of the 2-AP and the relevant sequence are shown. In panel A the 2-AP residue is incorporated at position 1, while in panel B the fluorescent base is at position 2. The spectra of the samples without protein or after incubation with wt UVDE or Y358A are in black as indicated. In grey is the spectrum recorded in the presence of Y358A and acrylamide.

### Mutant Y358A does not repair UV lesions *in vivo*.

The surprising increased incision activity of the Y358A mutant on the CPD and (6-4)PP lesions prompted us to test the ability of the mutant protein to repair UV-induced damage *in vivo*. For this purpose we used an *E. coli* strain lacking the *uvrA* and *uvrB* genes, which as a consequence of the absence of nucleotide excision repair (NER) is very UV sensitive (Figure 9, row 1). Introduction of a plasmid with the wt UVDE gene under control of the Plac promoter restores UV resistance (Figure 9, rows 2 and 6). The repair of UV lesions with only the UVDE protein is almost equal to the repair in a wt *E. coli* strain by the NER proteins (compare rows 2 and 3). Apparently the UVDE not only incises the DNA adjacent to the UV-induced lesions in *E. coli*, but the bacterium can effectively process the subsequent nicks to complete repair. The presence of the NER proteins together with UVDE even further improves the UV resistance (row 4). Apparently there are lesions left by the NER system (most likely CPD's) that are repaired by UVDE. This explains why some bacterial species (like *Bacillus subtilis*), which do possess an active NER machinery have also acquired the UVDE gene (Goosen and Moolenaar, 2008). Expression of the Y358A mutant protein in a NER-deficient background, however, only gave rise to very low UV survival (Figure 9, row 7), although the wild type and mutant proteins were expressed to the same level in the cells (results not shown).

What might explain this apparent discrepancy between the *in vivo* and *in vitro* activities of the mutant protein? One explanation might be that the binding of the protein to CPD lesions, which is already reduced *in vitro* might be much further reduced *in vivo* where there is a large excess of undamaged DNA. When we tested the incision of the CPD and (6-4)PP fragments in the presence of an excess of undamaged DNA, however, the activity of the mutant protein remained



**Figure 9.** UV survival test.

A nucleotide excision repair deficient *E. coli* strain containing a vector plasmid without UVDE (rows 1 and 5), or plasmids expressing wild type (rows 2 and 6) or mutant Y358A proteins (row 7) were irradiated with the indicated dose of UV light. As a control also a wt *E. coli* strain (row 3) and the same strain expressing wtUVDE (row 4) were included in the test.

very high (not shown). Another difference between the *in vivo* and *in vitro* conditions is the pH. The cellular pH of *E. coli* (and *S. pombe*) is around pH 7.5 (Padan *et al.*, 1981). The optimal pH for the wild type enzyme, however, is pH 6.5 (Kaur *et al.*, 1999 and our lab (not shown)) and all experiments have been conducted in a buffer of this lower pH. When we tested the wild type UVDE at the more physiological pH of 7.5 the enzyme still fully incised the DNA containing a CPD or (6-4)PP at 5 nM of protein (Figure 6B). The mutant protein, however, did not show any activity at all under these conditions (Figure 6B), although the binding of the mutant protein did not alter when the pH was varied (Table 5). This means that under physiological conditions, the catalytic activity of the mutant protein is strongly reduced, explaining the impaired repair of UV damage *in vivo*. We conclude from these experiments that Tyr358 not only plays a role in stabilization of the protein-DNA complex, but it also shields the active site from the influence of a higher pH of the solvent.

## DISCUSSION

One of the intriguing questions in DNA repair is how the different repair enzymes discriminate between damaged- and undamaged DNA. In some repair systems only one specific type of damage is recognized. An example of such an enzyme is uracil glycosylase where the enzyme contains a very specific uracil binding-pocket, which occludes binding of the four other bases (Parikh *et*

*al.*, 1998 and 2000). Nucleotide excision repair (NER) is an example of a repair system with a very broad substrate specificity where multiple enzymes participate in the damage recognition process, each protein probing different characteristics of the DNA structure (Truglio *et al.*, 2006, Gillet *et al.*, 2006). UVDE has also shown to be more versatile in damage recognition, although in contrast to NER only one relatively small protein is responsible for both damage-specific binding and incision. In this paper we present evidence that UVDE uses different parameters to recognize a DNA damage. Initially UVDE seems to probe the DNA for its bendability. We show that not only a CPD, (6-4)PP or abasic site but also a single-strand nick or gap significantly increase the binding affinity of the protein, although some of these complexes do not result in incision. The similarity between UVDE and Endo IV predicts that in the UVDE-DNA complex the DNA will be bent by about 90°. Any alteration in the DNA that increases its flexibility is therefore expected to stabilize the protein-DNA contacts. Most likely for the same reason mismatches and small insertions can be a substrate for UVDE (Kaur *et al.*, 1999, Gillet *et al.*, 2006). The bendability of the DNA alone, however, is not enough. Subsequently the protein needs to present the scissile bond to the active site, which is buried inside the protein, requiring rotation of the adjacent base from the DNA helix. For the CPD and (6-4)PP this means that both covalently linked bases need to be extruded from the helix into a proposed pocket of the protein (Paspaleva *et al.*, 2007). Our results with the AP site in different sequence contexts suggest that also on an AP lesion the protein will flip two residues: the abasic site together with its neighbour. Whether this will be the base at the 3' or 5' side of the AP site depends on the nature of these bases. The most preferred AP-substrate appears to be the AP site with a pyrimidine at its 3' side. The combination of an abasic site with a purine at the 3' side does not seem to be bound at all and instead on such a sequence UVDE will shift one position, now targeting the AP site together with the 5' flanking base. Apparently the protein pocket in which the two adjacent bases are proposed to flip can accommodate a normal or distorted (in the CPD and (6-4)PP) pyrimidine at the 3' side but not a purine, probably as a result of its larger size. Also for the base flanking the AP site at the 5' side a pyrimidine is preferred, although productive complexes can be formed with a 5' flanking purine as well albeit with much lower efficiency. We have presented evidence that UVDE can also form productive complexes on only one rotated pyrimidine residue if this residue is 3' of a broken backbone caused by a nick or a gap. Possibly because of the broken backbone the rotation of the two residues is no longer coupled, allowing insertion of only one pyrimidine into the protein pocket. Again here no productive complexes are formed when a purine flanks the nick in the backbone. Taken together we propose that exclusion of larger bases from the proposed protein pocket provides a second damage discrimination step.

Our results are in agreement with the reported UVDE-mediated incision of DNA containing a platinum-GG crosslink (Avery *et al.*, 1999). On this DNA the incision did not occur immediately 5' to the Pt-GG adduct which can be expected since the crosslinked guanine residues would be too large to fit into the proposed protein pocket. Instead incision was made 2 nt further downstream, placing the two pyrimidine residues that flank the Pt-GG adduct into the pocket

of the protein. Surprisingly in the same paper it was reported that UVDE incises DNA with an AP site in the CXA context immediately 5' to the AP site. This is in contrast to our observation where in the same sequence context the enzyme makes the incision 5' to the adjacent C residue. An important difference between the two substrates is that in our DNA fragments the AP site consists of a tetrahydrofuran (i.e. a closed ribose ring) whereas in the substrate used by Avery *et al.* (1999) the AP site was generated by removal of a uracil with uracil DNA glycosylase, resulting in a ring-opened ribose. Possibly the open ring structure provides more conformational freedom allowing rotation of only the AP site without the 3' flanking purine. It should be noted, however, that on this CXA sequence with the ring-opened ribose incision is much less efficient (ten times lower compared to a CPD (Avery *et al.*, 1999).

A thymine glycol (TG) lesion adjacent to a pyrimidine is as efficiently bound by UVDE as a CPD or (6-4)PP and the protein induces similar conformational changes in the bottom strand for all three lesions. Yet the TG-containing DNA is only poorly incised. *In silico* (Miaskiewicz *et al.*, 1995) and structural studies (Kim and Choi, 1995) showed that TG is not a planar molecule and the methyl group is oriented perpendicularly with respect to the pyrimidine ring, making the TG more 'bulky' compared to a normal thymine. It is not very likely that this will completely prevent the TG from being inserted into the protein pocket, but its conformation might make it more difficult to properly position the 5' phosphodiester bond into the catalytic site.

The 2-AP fluorescence studies presented in this paper show that the two bases opposite the residues that are proposed to be flipped into the UVDE protein pocket are significantly destacked. In the crystal structure of *T. thermophilus* UVDE, Tyr358 is seen to point out of the proposed DNA binding groove and therefore it is very likely that upon binding to DNA this residue will penetrate the DNA helix. One possibility is that Tyr358 will wedge between the two residues opposite the damage thereby disrupting the base stack and stabilizing the DNA kink. As a consequence the 3' residue might be partly rotated from the helix, since it was shown to be partly accessible to the solvent. Alternatively both residues in the complementary strand might take up an extrahelical position with Tyr358 occupying the vacated space and other residues of UVDE completely shielding the 5' base. We regard this possibility as less likely, however, since it is hard to envisage how this shielding would occur, since the bases are expected to be flipped away from the protein. In either case, the situation must be different from the base-flipping in the non-damaged strand by Endo IV, since in this complex the base is fully extrahelical but it does not make any specific contacts with the enzyme.

In the absence of Tyr358 destacking of the bases in the non-damaged strand is no longer observed and as a consequence binding to the CPD lesion is reduced. On the (6-4)PP binding is not affected indicating that the contribution of Tyr358 to the stability of the UVDE-DNA complex depends on the type of damage. A (6-4)PP lesion induces a kink of 44° into the DNA (Kim and Choi, 1995) whereas CPD lesion is less kinked by about 30° (Park *et al.*, 2002). Possibly if the DNA is already more kinked by the lesion itself the putative role of Tyr358 in stabilizing the kink becomes less important. The overall structure of DNA containing an AP site

does not show significant kinking (Chen *et al.*, 2008), explaining that on the CXT fragment the binding of the mutant protein is even further reduced.

The incision activities of the Y358A mutant on both CPD and (6-4)PP significantly differed from the wild type protein. At pH 6.5 the mutant protein was more active than the wild type, whereas at pH 7.5 the complete opposite was found resulting in an active wild type and inactive mutant protein. Since the pH did not affect the binding of either protein this means that Tyr358 influences the catalytic step of the reaction in a pH-dependent way.

In the co-crystal structure of Endo IV with its abasic site substrate there is also a tyrosine residue (Tyr72) that inserts into the DNA helix filling the gap left by the flipped-out AP site and stacking on its 5' neighbour (Hosfield *et al.*, 1999). Mutation of this residue into alanine only partially reduced the DNA binding, but caused a more dramatic decrease in catalysis. It was proposed that the hydrophobic residue shields the active site from the solvent to prevent influence on the  $pK_a$  of the catalytic Glutamate or hydroxide nucleophile (Garcin *et al.*, 2008). Indeed the structure of the Y72A mutant revealed the presence of water molecules in close proximity of the scissile phosphodiester bond (Garcin *et al.*, 2008). Also in UVDE Tyr358 might fulfil such a protective function. Although our fluorescence data show that this residue is more in contact with the non-damaged strand, its hydrophobic nature might similarly prevent access of the solvent to the active site. The fluorescence data show that the 5' opposing residue becomes more solvent exposed in the absence of Tyr358 and water molecules might therefore penetrate further into the active site. The pH dependence of the catalysis by the Y358A mutant, however, indicates that the solvent is not likely to influence the first step in catalysis, deprotonation of the attacking water molecule, since this would be expected to be less favourable at lower pH. Instead the mutant protein shows a higher activity at lower pH (6.5). Therefore we propose that in the mutant the solvent influences the last step of catalysis, the protonation of the departing 3'-O of the ribose. Under physiological conditions (pH 7.5) this is negatively influenced by the solvent, making the shielding effect of Tyr358 essential for the activity of the enzyme.

In conclusion we propose that although UVDE is capable of repairing different types of damage it is specialized in the removal of UV lesions (CPD and (6-4)PP) and we present the following model for its damage recognition. Initially the protein will bind to flexible sites in the DNA that allow bending of the DNA helix. Bending will be aided by insertion of the Tyr358 into the helix. Subsequently the protein will try to flip two potentially damaged bases into a pocket of the protein, which is located at the bottom of the DNA binding groove. Large bases will be excluded from this pocket, allowing entrance only of two pyrimidines or a pyrimidine together with an AP site. If insertion of these bases into the pocket is successful, the phosphodiester bond is brought in close vicinity of the metal ions in the catalytic site. The aromatic side chain of Tyr358 will shield the active site from the solvent allowing efficient catalysis at physiological pH.

## REFERENCES

- Allan BW and Reich NO, Targeted base stacking disruption by the *EcoRI* DNA methyltransferase, *Biochemistry* **35** (1996), pp. 14757-14762
- Aller P, Rould MA, Hogg M, Wallace SS and Doublet S, A structural rationale for stalling of a replicative DNA polymerase at the most common oxidative thymine lesion, thymine glycol, *Proc. Natl. Acad. Sci. USA*. **104** (2007), pp. 814-818
- Avery A, Kaur B, Taylor JS, Mello JA, Essigmann JM and Doetsch PW, Substrate specificity of ultraviolet DNA endonuclease (UVDE/Uve1p) from *Schizosaccharomyces pombe*, *Nucleic Acids Res.* **27** (1999), pp. 2256-2264
- Bandwar RP and Patel SS, Peculiar 2-aminopurine fluorescence monitors the dynamics of open complex formation by bacteriophage T7 RNA polymerase, *J. Biol. Chem.* **276** (2001), pp. 14075-14082
- Bowman K, Cidik K, Smith CA, Taylor JS, Doetsch PW and Freyer GA, A new ATP-independent DNA endonuclease from *Schizosaccharomyces pombe* that recognizes cyclobutane pyrimidine dimers and 6-4 photoproducts, *Nucleic Acids Res.* **22** (1994), pp. 3026-3032
- Chen J, Dupradeau F-Y, Case DA, Turner CJ and Stubbe JA, DNA oligonucleotides with A, T, G or C opposite an abasic site: structure and dynamics, *Nucleic Acids Res.* **36** (2008), pp. 253-262
- Christine KS, MacFarlane AW, Yang K and Stanley RJ, Cyclobutylpyrimidine dimer base flipping by DNA photolyase, *J. Biol. Chem.* **277** (2002), pp. 38339-38344
- Earl AM, Rankin SK, Kim KP, Lamendola ON and Battista J, Genetic evidence that the *uvrE* gene product of *Deinococcus radiodurans* R1 is a UV damage endonuclease, *J. Bacteriol.* **184** (2002), pp. 1003-1009
- Garcin ED, Hosfield DJ, Desai SA, Haas BJ, Björas M, Cunningham RP and Tainer JA, DNA apurinic-apyrimidinic site binding and excision by endonuclease IV, *Nature Struct. & Mol. Biol.* **15** (2008), pp. 515-522
- Gillet LC and Schärer OD, Molecular mechanisms of mammalian global genome nucleotide excision repair, *Chem. Rev.* **106** (2006), pp. 253-276
- Goosen N and Moolenaar GF, Repair of UV damage in bacteria, *DNA Repair (Amst.)* **7** (2008), pp. 353-379
- Holz B, Klimasaukas S, Serva S and Weinhold E, 2-aminopurine as a fluorescent probe for DNA base flipping by methyltransferases, *Nucleic Acids Res.* **26** (1998), pp. 1076-1083
- Hosfield DJ, Guan Y, Haas BJ, Cunningham RP and Tainer JA, Structure of the DNA repair enzyme Endo IV and its DNA complex: double-nucleotide flipping at abasic sites and three-metal-ion catalysis, *Cell* **98** (1999), pp. 397-408
- Iwai S, Chemical synthesis of oligonucleotides containing damaged bases for biological studies, *Nucleosides Nucleotides Nucleic Acids* **25** (2006), pp. 561-582
- Kanno S, Iwai S, Takao M and Yasui A, Repair of apurinic/apyrimidinic sites by UV damage endonuclease; a repair protein for UV and oxidative damage, *Nucleic Acids Res.* **27** (1999), pp. 3096-3103
- Kaur B, Fraser LA, Freyer GA, Davey S and Doetsch PW, A Uve1p-mediated mismatch repair pathway in *Schizosaccharomyces pombe*, *Mol. Cell. Biol.* **19** (1999), pp. 4703-4710
- Kim J-K and Choi B, The solution structure of DNA duplex-decamer containing the (6-4) photoproduct of thymidylyl(3'→5')thymidine by NMR and relaxation matrix refinement, *Eur. J. Biochem.* **228** (1995), pp. 849-854
- Malta E, Moolenaar GF and Goosen N, Base flipping in nucleotide excision repair, *J. Biol. Chem.* **281** (2006), pp. 2184-2194

- Miaskiewicz K, Miller J, Ornstein R and Osman R, Molecular dynamics simulations of the effects of ring-saturated thymine lesions on DNA structure, *Biopolymers* **35** (1995), pp. 113-124
- McCullough AK, Dodson ML, Schärer OD and Lloyd RS, The role of base flipping in damage recognition and catalysis by T4 endonuclease V, *J. Biol. Chem.* **272** (1997), pp. 27210-27217
- Moolenaar GF, Schut M and Goosen N, Binding of the UvrB dimer to non-damaged and damaged DNA: residues Y92 and Y93 influence the stability of both subunits, *DNA Repair (Amst.)* **4** (2005), pp. 699-713
- Padan I, Zilberstein D and Schuldiner S, pH homeostasis in bacteria, *Biochim. Biophys. Acta* **650** (1981), pp. 151-166
- Parikh SS, Mol CD, Slupphaug G, Bharati S, Krokan HE and Tainer JA, Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA, *EMBO J.* **17** (1998) pp. 5214-5226
- Parikh SS, Putnam CD and Tainer JA, Lessons learned from structural results on uracil-DNA glycosylase, *Mutat. Res.* **460** (2000) pp. 183-199
- Park H, Zhang K, Ren Y, Nadji S, Sinha N, Taylor J and Kang C, Crystal structure of a DNA decamer containing a cis-syn thymine dimer, *Proc. Natl. Acad. Sci. USA.* **99** (2002), pp. 15965-15970
- Paspaleva K, Thomassen E, Pannu NS, Iwai S, Moolenaar GF, Goosen N and Abrahams JP, Crystal structure of the DNA repair enzyme UV damage endonuclease, *Structure* **15** (2007), pp. 1316-1324
- Rai P, Cole TD, Thompson E, Millar DP and Linn S, Steady-state and time-resolved fluorescence studies indicate an unusual conformation of 2-aminopurine within ATAT and TATA duplex DNA sequences, *Nucleic Acids Res.* **31** (2003), pp. 2323-2332
- Studier FW, Rosenberg AH, Dunn JJ and Dubendorff JW, Use of T7 RNA polymerase to direct expression of cloned genes, *Methods Enzymol.* **185** (1990), pp. 60-89
- Takao M, Yonemasu R, Yamamoto K and Yasui A, Characterization of a UV endonuclease gene from the fission yeast *Schizosaccharomyces pombe* and its bacterial homolog, *Nucleic Acids Res.* **24** (1996), pp. 1267-1271
- Truglio JJ, Croteau DL, Van Houten B and Kisker C, Prokaryotic nucleotide excision repair, the UvrABC system, *Chem. Rev.* **106** (2006) pp. 233-252
- Verhoeven EE, van Kesteren M, Turner JJ, van der Marel GA, van Boom JH, Moolenaar GF and Goosen N, The C-terminal region of *Escherichia coli* UvrC contributes to the flexibility of the UvrABC nucleotide excision repair system, *Nucleic Acids Res.* **30** (2002), pp. 2492-2500
- Yajima H, Takao M, Yasuhira S, Zhao JH, Ishii C, Inoue H and Yasui A, A eukaryotic gene encoding an endonuclease that specifically repairs DNA damaged by ultraviolet light, *EMBO J.* **14** (1995), pp. 2393-2399

# Active site organisation of UVDE - a $Mn^{2+}$ dependent nuclease

Chapter

5

Keti Paspaleva, Geri F. Moolenaar, Nora Goosen

---

Laboratory of molecular genetics, Leiden Institute of Chemistry, Gorlaeus Laboratories,  
Leiden University, Einsteinweg 55, 2300 RA Leiden, The Netherlands

## SUMMARY

UV damage endonuclease (UVDE) initiates an alternative repair pathway for UV lesions by introducing a nick immediately 5' to both cyclobutane pyrimidine dimers and 6-4 photoproducts. Discovered for the first time in *Schizosaccharomyces pombe*, a broad substrate specificity has been proposed for this nuclease, including not only UV lesions, but also abasic sites, single strand nicks and thymine glycols. By comparing the activity of UVDE proteins from *Schizosaccharomyces pombe* (Spo), *Bacillus subtilis* (Bsu) and *Thermus thermophilus* (Tth) we show that the broad substrate specificity is restrained to the eukaryotic Spo homologue, while the Tth-UVDE functions as an UV damage specific repair enzyme. In this paper we give clear evidence that  $Mn^{2+}$  is the catalytic cofactor for all three UVDE proteins. We propose that full occupation of the UVDE metal binding site occurs only after DNA binding. We show that for the Spo- and the Bsu-UVDE  $Mg^{2+}$  has a specific role in reducing the non-specific DNA binding. In contrast, for the Tth enzyme  $Mg^{2+}$  is shown to inhibit the protein activity by competing with  $Mn^{2+}$  in the active site. We further show that the inability of Bsu-UVDE to incise DNA containing a thymine glycol or a single strand nick is not due to lack of damage recognition. Instead, after binding to these lesions the Bsu protein fails to induce the required destacking of the bases opposite the damage.

## INTRODUCTION

Ultraviolet light induces two major types of DNA damage: cyclobutane pyrimidine dimers (CPDs) and (6-4) photoproducts (6-4PP). To ensure cell survival, various DNA repair pathways have evolved to remove these lesions. One effective mechanism is the UV damage endonuclease (UVDE), where a single UVDE enzyme recognises and subsequently nicks DNA containing a CPD or a 6-4PP. UVDE was identified for the first time in the fission yeast *Schizosaccharomyces pombe* (Bowman *et al.*, 1994). Initially broad substrate specificity has been suggested for this nuclease, including not only UV lesions, but also abasic sites and some nucleotide mismatches (Kanno *et al.*, 1999), (Kaur *et al.*, 1999). Later studies, however, revealed that the *S. pombe* UVDE recognises the abasic site in a sequence dependent manner, with a very low activity when the damage is flanked by two purines (Paspaleva *et al.*, 2009).

UVDE orthologues have been discovered in a number of fungi and eubacterial species, as well as in four archaeobacteria (Goosen and Moolenaar, 2008). On the amino acid level there are some intriguing differences between the pro- and eukaryotic proteins, since the bacterial UVDEs are notably shorter and lack the N-terminal 200 amino acids. The C-terminal part of the protein is also not conserved and varies in length (Goosen and Moolenaar, 2008).

The structure of *T. thermophilus* UVDE, solved with 1.55 Å resolution, revealed significant similarities with the BER enzyme Endonuclease IV, both being TIM-barrel metal dependent nucleases (Paspaleva *et al.*, 2007). It is currently believed that the catalytic mechanism of UVDE might be very similar to the one of Endo IV, where the DNA is cleaved with the help of trimetal Zn<sup>2+</sup> cluster (Hossfield *et al.*, 1999).

In the UVDE structure the catalytic site was also seen to have the potential to accommodate three ions, although due to low occupation, the anomalous density maps failed to elucidate their nature. Crystallisation in the presence of Mn<sup>2+</sup> ions yielded structure with improved occupancies of the active site metals, highly suggesting that Mn<sup>2+</sup> is able to bind within the enzyme catalytic site (Meulenbroek *et al.*, 2009). Biochemical tests further supported this hypothesis, since Mn<sup>2+</sup> was seen to be requirement for the *T. thermophilus* UVDE catalytic activity (Paspaleva *et al.*, 2007). The UVDE active site metals superimpose on the three zinc sites of Endo IV (Paspaleva *et al.*, 2007), implying that Mn1 and/or Mn2 may be involved in creating the water derived nucleophile.

Interestingly, although the crystal structure of Endo IV reveals the presence of three Zn<sup>2+</sup> ions, previous data suggest that for its function it can use other divalent metals such as Mn<sup>2+</sup> (Levin *et al.*, 1988), Co<sup>2+</sup> (Levin *et al.*, 1991) and Mg<sup>2+</sup> (Kutyavin *et al.*, 2006). Although Mn<sup>2+</sup> was proposed as a possible cofactor for the *T. thermophilus* UVDE, little is known about the ionic requirements of UVDE in other organisms. The vast majority of biochemical tests with the *S. pombe* homologue have been performed with a combination of 1 mM Mn<sup>2+</sup> and 10 mM Mg<sup>2+</sup> in the reaction mix (reviewed in Doetch *et al.*, 2006). In some studies (Kanno *et al.*, 1999) only Mg<sup>2+</sup> was utilized, showing that the *S. pombe* protein can use this metal ion as a catalytic

cofactor. For the *T. thermophilus* UVDE, however, Mg<sup>2+</sup> is not able to stimulate the enzymatic activity (Paspaleva *et al.*, 2007) and the UVDE from *Deinococcus radiodurans* was also reported to specifically require Mn<sup>2+</sup> (Evans and Moseley, 1985).

Here we give evidence that Mn<sup>2+</sup> is the optimal cofactor for the UVDE enzymes from *S. pombe*, *B. subtilis* and *T. thermophilus*. Other metals such as Co<sup>2+</sup> and Ca<sup>2+</sup> could substitute Mn<sup>2+</sup>, but to a lesser extent and depending on the type of protein and DNA lesion. Furthermore, we present evidence that the UVDE active site gets fully occupied only after DNA binding. We also compared the activity of the three different UVDE enzymes not only on UV lesions but also on DNA substrates containing an abasic site, a single nucleotide nick or a thymine glycol, showing that the previously reported broad substrate specificity is restricted to the eukaryotic *S. pombe* enzyme.

## MATERIALS AND METHODS

### Proteins

Plasmid pETUVDEΔ228, encoding residues 229 - 599 of *S. pombe* UVDE, fused to an N-terminal 10 × His tag has been described before (Paspaleva *et al.*, 2009). All *S. pombe* UVDE point mutations were constructed by PCR and verified by sequencing for the absence of additional PCR-induced mutations. The mutant UVDE proteins were purified using the same procedure as described for the wild type enzyme (Paspaleva *et al.*, 2009) and they all showed the same elution/purification profile.

The *T. thermophilus* UVDE protein was purified as described before (Paspaleva *et al.*, 2007).

The *B. subtilis* UVDE gene was amplified by PCR, using primers fusing the C-terminal part of the protein to a 10 × His tag. Subsequently, the gene was inserted into the *Nde*I and *Bam*HI restriction sites of pET11a. The resulting *B. subtilis* UVDE expression vector (pUD17) was transformed into *E. coli* BL21/codon+ (Studier *et al.*, 1990). The UVDE protein was purified from cells of a 2 l culture, harvested 2 h after induction by IPTG and lysed by sonication in 6 ml lysis buffer (50 mM Tris pH 7.5, 150 mM NaCl, 10 mM β-mercaptoethanol, 10 % glycerol, 1 % Triton X-100, 1 % sarcosyl). The lysate was loaded on a HiTrap-chelating Ni column, which was equilibrated with 20 mM Tris pH 7.5. The protein was eluted with a 20 mM to 250 mM imidazole gradient in 20 mM Tris pH 7.5. The UVDE containing fractions were loaded on a hydroxyapatite column, equilibrated with 10 mM KPO<sub>4</sub> (pH 6.5). The protein was eluted with a gradient of 200 mM to 400 mM KPO<sub>4</sub> (pH 6.5). The pooled fractions were subsequently loaded on a P11 phosphocellulose column, equilibrated with 300 mM KPO<sub>4</sub> (pH 6.5) and eluted with 0 M to 1 M gradient of NaCl in the same buffer. Finally, the purified UVDE protein fractions were applied to a Nap<sup>TM</sup>5 gel filtration column (GE Healthcare) and eluted in 20 mM Tris pH 7.5, 150 mM NaCl and 10 % glycerol.

### UV survival test

The *Xba*I - *Bam*HI fragments from pETUVDE $\Delta$ 228 (*S. pombe* UVDE) and pUD17 (*B. subtilis* UVDE) were inserted in pIC-19R, resulting in the expression of the wild type and mutant UVDE proteins under the control of the Plac promoter. The resulting pIC-19R derivatives were introduced in CS 5018 ( $\Delta$ *uvrA*,  $\Delta$ *uvrB*) (Moolenaar *et al.*, 1994). Cells were grown in LB including 40  $\mu$ g/ml ampicillin and 1 mM IPTG until OD<sub>600</sub> of 0.3. The cultures were then diluted ten times and 1  $\mu$ l drops were spotted on LB plates containing 40  $\mu$ g/ml ampicillin and 1 mM IPTG. The drops were subsequently irradiated with the indicated dose of UV light and the plates were incubated overnight at 30°C.

### DNA substrates

The 30 bp DNA fragments (5'-CTCGTCAGCATCTTCATCATAACAGTC AGTG-3') with TT representing the position of the UV lesion (CPD or 6-4PP) were synthesised as described (Iwai, 2006). The same DNA sequence 5'-CTCGTCAGCATCXTCATCATAACAGTCAGTG-3', with X representing the position of the damage was used for the abasic site (via incorporation of a tetrahydrofuran-dSpacer) or the thymine glycol and obtained commercially (Eurogentec, Belgium). For the studies performed with the ssDNA nick a 50 bp fragment was used 5'-CGTGTGAGGTCGTTCTGAGGTTT\*TTTTGTAATGTGCCCGTAAGTAATCCC-3', where the star represents the position of the nick. All substrates containing 2-aminopurine (2 - AP) have been described (Paspaleva *et al.*, 2009).

The DNA fragments were 5' radioactively labelled using polynucleotide kinase as described (Verhoeven *et al.*, 2002).

### Incision assay

The 5' terminally labelled DNA substrates (0.2 nM) were incubated with the indicated amounts of UVDE in a 20  $\mu$ l reaction mix (20 mM HEPES pH 6.5, 100 mM NaCl) in the presence of the shown amount of divalent metal ions. For some experiments (as specified) the 20 mM HEPES pH 6.5 was substituted with 20 mM Tris with the same pH. After 15 minutes incubation at 30°C (*S. pombe* and *B. subtilis* UVDE) or 55°C (*T. thermophilus* UVDE), the reaction was terminated by adding 3  $\mu$ l of 0.33 M EDTA, 3.3 % SDS and 2.4  $\mu$ l glycogen (4  $\mu$ g /  $\mu$ l), followed by an ethanol precipitation. The incision products were visualised on a 15 % acrylamide gel.

### Filter Binding Assay

The filter binding assays were performed in 20  $\mu$ l samples containing 5 nM or 50 nM UVDE and 4 nM of the terminally labelled DNA substrates in a reaction buffer containing 20 mM Tris pH 6.5, 100 mM NaCl. The mixes were incubated for 10 min at 30°C (*B. subtilis* and *S. pombe* UVDE) or 55°C (*T. thermophilus* UVDE). If needed, metal cofactors (1 mM MnCl<sub>2</sub> and/or 10 mM MgCl<sub>2</sub>) were included in the reaction mix. At the end of the incubation time 0.5 ml of preheated (30°C or 55°C) reaction buffer was added to each sample. The mixture was

subsequently poured over a nitrocellulose filter (Millipore 0.45 µm HA) and the filter was rinsed three times with 0.5 ml of the reaction buffer. Each sample was corrected for the amount of DNA retained on a filter in the absence of protein. Binding is expressed as the percentage of input DNA retained on the filter by the enzymes. For determination of the UVDE binding activity under the conditions used in the 2-AP measurements 4 nM of the terminally labelled substrates were mixed with 0.5 µM unlabeled DNA of the same construct and incubated with *S. pombe* UVDE (2.5 µM) and *B. subtilis* UVDE (2.5 µM or 4 µM) for 10 minutes at 30°C.

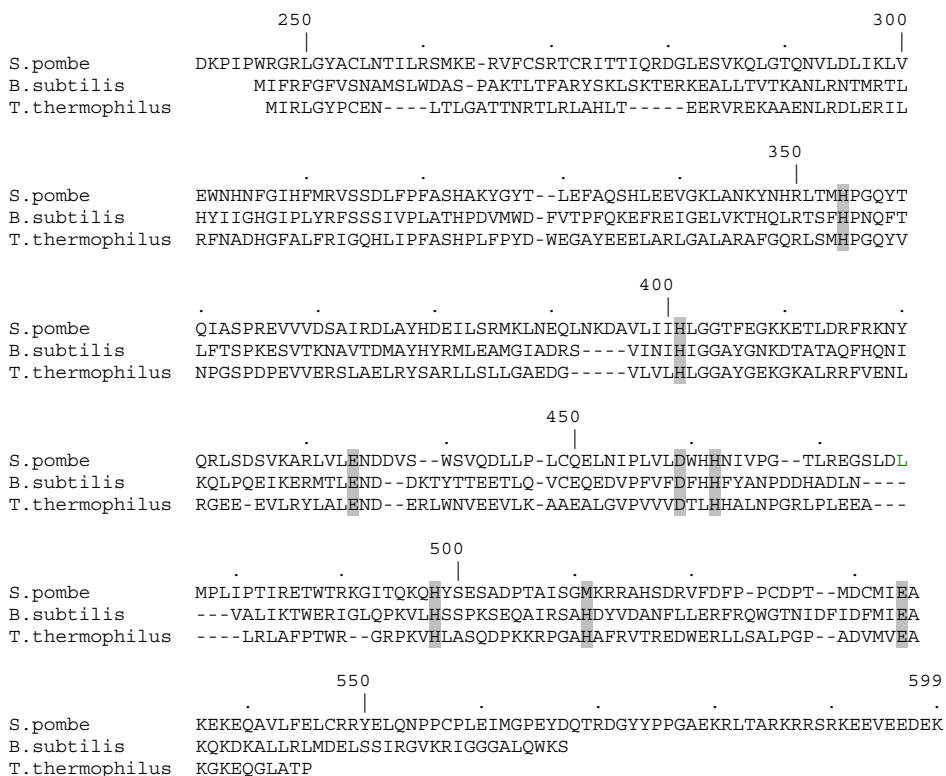
### 2-AP measurements

The 2-AP measurements were conducted in 60-µl samples as described before (Paspaleva *et al.*, 2009), using 0.5 µM DNA, 2.5 µM *S. pombe* UVDE and 2.5 µM or 4 µM *B. subtilis* UVDE. The samples were transferred to a 3 mm x 3 mm quartz cuvette and placed in the fluorimeter. Fluorescence emission spectra were obtained using a PerkinElmer LS 50B fluorimeter, connected to a temperature variable water bath to maintain a temperature inside the cuvette of 30°C. The excitation wavelength was set at 310 nm, and the emission spectra were obtained by scanning from 325 to 475-nm.

## RESULTS

### Mn<sup>2+</sup> is the optimal UVDE catalytic cofactor for all three enzymes

The similarities between the structures of UVDE from *Thermus thermophilus* (Tth-UVDE) and Endo IV suggest a comparable catalytic mechanism and likely utilization of three metal ion cluster for the DNA cleavage. Although Mn<sup>2+</sup> was previously shown to be able to bind within the active site of Tth-UVDE and promote catalysis (Paspaleva *et al.*, 2007; Meulenbroek *et al.*, 2009) it still remains unclear if it is the optimal UVDE cofactor. Furthermore, it is still an open question whether the UVDE proteins from different organisms share the same divalent metal requirements, although as seen in Figure 1 the metal coordinating residues are highly conserved. To obtain a detailed overview of the enzyme cofactor requirements we tested the incision efficiencies of UVDE proteins from *Schizosaccharomyces pombe* (Spo-UVDE), *Bacillus subtilis* (Bsu-UVDE) and *T. thermophilus* (Tth-UVDE) on DNA fragments containing a CPD, a 6-4PP or an abasic site in the presence of different divalent ions (Table 1). The abasic site was chosen to be flanked by two pyrimidines, since it has been shown that the incision activity of Spo-UVDE on this lesion strongly depends on the presence of a neighbouring pyrimidine residue (Paspaleva *et al.*, 2009).



**Figure 1:** Alignment of the UVDE homologues.

The amino acid sequence of UVDE from *S. pombe* (Takao *et al.*, 1996) is aligned with UVDE from *T. thermophilus* (Henne *et al.*, 2004) and *B. subtilis* (Kunst *et al.*, 1997). The metal coordinating residues are highlighted in grey. The numbering corresponds to the amino acids position in the *S. pombe* UVDE.

As seen in Table 1 (A and B), when  $Mn^{2+}$  is added to the reaction mixtures, Spo-UVDE and Bsu-UVDE exhibit similar high activities on all tested DNA damages. In the absence of  $Mn^{2+}$  both nucleases do not show any catalytic activity, pointing to the fact that their active sites are not fully occupied with three metal ions. Under this condition the complex formation of both enzymes is comparable, being equally high on the CPD, the 6-4PP and the abasic site (Table 2, A and B). Notably, the amount of protein-DNA complexes is not influenced by the presence of  $Mn^{2+}$ , since the binding efficiencies of both proteins are similar in the absence of additional metals or in the presence of  $Mn^{2+}$  (Table 2, A and B). This fact highly suggests that full occupation of the UVDE active site is not a requirement for stable DNA binding.

Tth-UVDE shows high incision on substrates containing CPD or 6-4PP, but as previously illustrated (Paspaleva *et al.*, 2007), cuts the abasic site containing DNA only 20 % (Table 1C). Interestingly, in the absence of metal ions or in the presence of  $Mn^{2+}$  Tth-UVDE distinguishes itself

**Table 1:** Activity of UVDE enzymes in the presence of different divalent metal ions.

A							
Organism:	-	Mn <sup>2+</sup>	Co <sup>2+</sup>	Mg <sup>2+</sup>	Ca <sup>2+</sup>	Zn <sup>2+</sup>	Cu <sup>2+</sup>
<i>S. pombe</i>							
CPD	< 1	97 ± 2	94 ± 2	< 1	5 ± 1	5 ± 2	2
6-4PP	< 1	93 ± 4	79 ± 1	1	20 ± 2	5 ± 1	1
Ap site	< 1	98 ± 1	10 ± 1	< 1	< 1	< 1	< 1
B							
Organism:	-	Mn <sup>2+</sup>	Co <sup>2+</sup>	Mg <sup>2+</sup>	Ca <sup>2+</sup>	Zn <sup>2+</sup>	Cu <sup>2+</sup>
<i>B. subtilis</i>							
CPD	< 1	99 ± 1	20 ± 2	< 1	< 1	5 ± 2	10 ± 3
6-4PP	< 1	94 ± 2	50 ± 3	< 1	< 1	< 1	< 1
Ap site	< 1	97 ± 1	2	< 1	< 1	< 1	< 1
C							
Organism:	-	Mn <sup>2+</sup>	Co <sup>2+</sup>	Mg <sup>2+</sup>	Ca <sup>2+</sup>	Zn <sup>2+</sup>	Cu <sup>2+</sup>
<i>T. thermophilus</i>							
CPD	< 1	94 ± 1	88 ± 4	< 1	< 1	< 1	< 1
6-4PP	< 1	91 ± 1	50 ± 2	< 1	< 1	< 1	< 1
Ap site	< 1	20 ± 2	20 ± 1	< 1	< 1	< 1	< 1

Incubation was done with 5 nM UVDE and 0.2 nM DNA in the presence of 1 mM of the indicated divalent metal ions or in the absence of additional metals (indicated as -).

as a weak DNA binder, since only 34 % of complex formation is detected on the CPD lesion and about 20 % on the 6-4PP and the abasic site. Even more, at 5 nM protein concentration, which was sufficient to promote 50 % – 60 % binding for Bsu-UVDE and Spo-UVDE, no binding of the Tth-UVDE can be detected at all. Although weak, the Tth-UVDE complex formation is still damage specific, since the complex formation on the undamaged DNA is significantly lower (Table 2C). A probable explanation might be the temperature optimum of the thermophilic enzyme, since it is plausible that at higher temperatures the protein-DNA contacts are weaker and easily disrupted during the experimental procedure. Catalytically the Tth-UVDE is as active as the Bsu- and the Spo-UVDE proteins (Figure 3C), suggesting that transient binding is sufficient for a high incision level. Although the incision of the abasic site by Tth-UVDE is very low, the binding to this substrate is comparable to that observed on the 6-4PP (Table 2C). Apparently both types of damages are equally recognised by the Tth-UVDE enzyme, but subsequent incision cannot efficiently take place at the abasic site.

As seen in Table 1, the highest incision for all tested UVDE proteins is observed in the presence of Mn<sup>2+</sup>, indicating that it is the optimal UVDE cofactor. Co<sup>2+</sup> can partially substitute Mn<sup>2+</sup>, but depending on the lesion and with varying efficiencies for the different UVDE enzymes. For Spo-UVDE Co<sup>2+</sup> can fully promote activity for the CPD incision, although studies performed with lower protein concentration suggest that Mn<sup>2+</sup> is slightly more beneficial (data not shown). On the 6-4PP Co<sup>2+</sup> appears a somewhat less efficient cofactor and on the abasic

**Table 2:** Binding activity of UVDE enzymes in the presence or absence of additional divalent metals.

## A. Spo UVDE

Cofactors:	CPD		6-4PP		Abasic site		Undam.DNA	
	50 nM	5 nM	50 nM	5 nM	50 nM	5 nM	50 nM	5 nM
-	91 ± 1	55 ± 3	92 ± 1	57 ± 1	90 ± 1	53 ± 3	9 ± 1	3 ± 1
Mn <sup>2+</sup>	94 ± 2	53 ± 2	93 ± 1	60 ± 1	93 ± 2	58 ± 1	23 ± 1	6 ± 1
Mg <sup>2+</sup>	90 ± 2	50 ± 4	91 ± 2	57 ± 1	85 ± 4	57 ± 2	3 ± 1	1 ± 2
Mn <sup>2+</sup> + Mg <sup>2+</sup>	94 ± 1	62 ± 2	95 ± 1	66 ± 1	94 ± 1	65 ± 1	11 ± 1	2 ± 2

## B. Bsu UVDE

Cofactors:	CPD		6-4PP		Abasic site		Undam.DNA	
	50 nM	5 nM	50 nM	5 nM	50 nM	5 nM	50 nM	5 nM
-	81 ± 2	55 ± 2	86 ± 1	56 ± 3	82 ± 2	60 ± 1	19 ± 2	3 ± 2
Mn <sup>2+</sup>	80 ± 3	56 ± 2	82 ± 1	59 ± 1	84 ± 2	59 ± 1	40 ± 2	6 ± 2
Mg <sup>2+</sup>	77 ± 1	51 ± 2	84 ± 1	58 ± 3	79 ± 1	58 ± 3	7 ± 1	3 ± 1
Mn <sup>2+</sup> + Mg <sup>2+</sup>	82 ± 1	61 ± 3	87 ± 1	67 ± 1	86 ± 1	66 ± 1	21 ± 3	5 ± 2

## C. Tth UVDE

Cofactors:	CPD		6-4PP		Abasic site		Undam.DNA	
	50 nM	5 nM	50 nM	5 nM	50 nM	5 nM	50 nM	5 nM
-	34 ± 2	2 ± 1	19 ± 3	< 1	20 ± 2	1 ± 1	3 ± 1	1 ± 1
Mn <sup>2+</sup>	33 ± 2	1 ± 1	20 ± 1	< 1	23 ± 1	< 1	3 ± 1	1 ± 1
Mg <sup>2+</sup>	3 ± 1	< 1	4 ± 1	< 1	4 ± 1	< 1	2 ± 1	< 1
Mn <sup>2+</sup> + Mg <sup>2+</sup>	4 ± 1	< 1	21 ± 1	< 1	20 ± 1	< 1	3 ± 1	< 1

Incubation was done with 4 nM DNA and 50 nM or 5 nM UVDE in the absence of divalent metal ions or in the presence of 1 mM Mn<sup>2+</sup> and/or 10 mM Mg<sup>2+</sup> as indicated. The binding is expressed as the percentage of total input DNA.

site its activity is severely reduced (Table 1A). This observation suggests that the type of lesion strongly influences the productive coordination of the divalent metal ions.

Notably, for the Bsu-UVDE Co<sup>2+</sup> is less capable of replacing Mn<sup>2+</sup> for its catalytic function on all three substrates. It gives rise to only 20 % incision on the CPD, 50 % on the 6-4PP and hardly any nicking (2 %) on the abasic site (Table 1B). The reduction in the incision efficiency is not due to loss of protein-DNA complexes, since the binding of Bsu-UVDE to damaged DNA is similar in the presence of Co<sup>2+</sup> or Mn<sup>2+</sup> (data not shown).

An interesting observation is that for Bsu-UVDE Co<sup>2+</sup> promotes higher activity on the 6-4PP in comparison with the CPD, while for the Spo enzyme the opposite is observed. Interestingly, Co<sup>2+</sup> and Mn<sup>2+</sup> serve as equally good cofactors for the Tth-UVDE incision on the abasic site. Apparently, although the metal coordinating residues are fully conserved, the active sites of Spo-UVDE, Tth-UVDE and Bsu-UVDE might be structurally different.

We also tested a number of other divalent metals for their ability to support the UVDE incisions. As seen in Table 1 Mg<sup>2+</sup> is not able to promote any detectable incision for all three UVDE enzymes. Both Bsu-UVDE and Spo-UVDE can utilize Zn<sup>2+</sup> and Cu<sup>2+</sup>, albeit with very

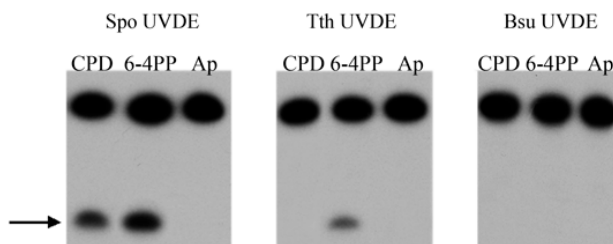
low efficiency. The Spo protein is the only one from the tested UVDEs, which is able to use  $Ca^{2+}$  as a cofactor, but only on the 6-4PP (20 % incision). Bearing in mind the large ionic radius (100 pm) of  $Ca^{2+}$  (all other tested divalent metals have ionic radii in the 70 pm range) this shows that Spo-UVDE possesses a remarkable flexibility in its active site. Unlike the other homologues Tth-UVDE is functional only in the presence of  $Mn^{2+}$  and  $Co^{2+}$ . The reduction of the incision levels is not due to disturbance in the protein-DNA complexes, since the binding of all tested UVDEs in the presence of  $Mn^{2+}$ ,  $Co^{2+}$  or  $Ca^{2+}$  is comparable (data not shown).

Taken together, our data clearly demonstrate that for all UVDE proteins the presence of  $Mn^{2+}$  is the most beneficial condition in terms of incision efficiency. Other divalent metals, such as  $Co^{2+}$  and  $Ca^{2+}$  can partially substitute  $Mn^{2+}$ , the level of substitution being highly dependent on the DNA lesion and the type of enzyme suggesting structural differences in the active sites of the different UVDE proteins.

### **$Mg^{2+}$ prevents unspecific binding of Spo- and Bsu-UVDE**

Although for many other nucleases it has been shown that  $Mn^{2+}$  and  $Mg^{2+}$  can substitute each other in function, this is not the case for UVDE. Since no detectable incision for all tested UVDE enzymes was observed with 1 mM  $Mg^{2+}$  (Table 1) we also used higher (10 mM)  $Mg^{2+}$  concentrations. Under these conditions very low incision is obtained with Spo-UVDE, but only on the UV lesions (Figure 2). Bsu-UVDE doesn't show any nicking and the Tth-UVDE exhibits an extremely low activity on the 6-4PP and no activity on the CPD (Figure 2).

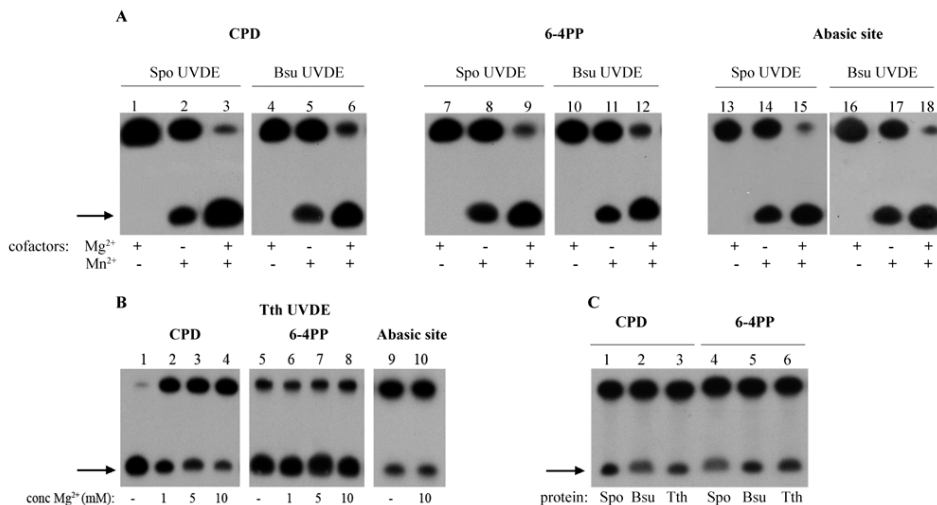
In literature most of the biochemical tests with Spo-UVDE have been performed in the presence of both  $Mn^{2+}$  and  $Mg^{2+}$  ions suggesting optimal activity under these conditions (reviewed in Doetch *et al.*, 2006). Indeed, when we used a lower protein concentration (2 nM), incision on all three substrates by the Spo enzyme is higher in the presence of both 1 mM  $Mn^{2+}$  and 10 mM  $Mg^{2+}$  than with  $Mn^{2+}$  alone (Figure 3A). The same is observed for Bsu-UVDE (Figure 3A). One possible explanation for the observed stimulating effect of  $Mg^{2+}$  might be that it becomes part of a catalytically more potent mixed  $Mn^{2+}/Mg^{2+}$  active site. Alternatively its effect might not be



**Figure 2:** Effect of 10 mM  $Mg^{2+}$  on the incision efficiency of different UVDE enzymes.

5 nM of UVDE from *T. thermophilus* (Tth), *S. pombe* (Spo) or *B. subtilis* (Bsu) was incubated for 15 minutes with 0.2 nM DNA substrates containing a CPD, 6-4PP or an abasic site (Ap) lesion in the presence of 10 mM  $Mg^{2+}$ . The incision product is indicated with an arrow.

related to a direct binding within the metal coordination site, but to an influence on the stability of the protein-DNA complexes. To discriminate between these two possibilities we performed DNA binding studies in the presence or absence of additional  $Mg^{2+}$ .



**Figure 3:** Effect of  $Mn^{2+}$  and  $Mg^{2+}$  on the incision efficiency of different UVDE enzymes.

A. The UVDE enzymes (2 nM) from *S. pombe* (Spo) and *B. subtilis* (Bsu) were incubated with 0.2 nM DNA containing a CPD, a 6-4PP and abasic site lesions in the presence of 1 mM  $Mn^{2+}$  or/and 10 mM  $Mg^{2+}$  as indicated.

B. *T. thermophilus* (Tth) UVDE (5 nM) was incubated with the same DNA substrates in the presence of 1 mM  $Mn^{2+}$  and different concentration of  $Mg^{2+}$  ions as indicated.

C. Spo, Bsu and Tth UVDE (0.25 nM) were incubated for 15 minutes with 0.2 nM CPD and 6-4PP containing DNA fragments. The position of the incision products is shown with arrows.

As shown in Table 2, in the absence of divalent metals the binding of Bsu-UVDE and Spo-UVDE to CPD, 6-4PP and the abasic site is comparably high. Addition of  $Mg^{2+}$  or  $Mn^{2+}$  alone does not significantly change the amount of detected complexes on any of the tested lesions. With the combination of  $Mn^{2+}$  and  $Mg^{2+}$ , however, at low protein concentrations (5 nM), a slight stimulating effect is observed.

When we tested the binding of both proteins to undamaged DNA, much more pronounced effect of the metal ions is observed. For both enzymes  $Mn^{2+}$  stimulates the unspecific binding at 50 nM of protein (Table 2, A and B). This is an interesting observation, since the role of the metal cofactors is generally limited to the catalytic step and little is known about their influence on the DNA binding. Notably, Bsu-UVDE shows higher affinity for the undamaged DNA substrate than the Spo protein. On the damaged DNA substrates this  $Mn^{2+}$  effect is less visible, since it is probably masked by the much stronger impact of the DNA lesion on the enzyme-DNA

interactions. Addition of  $Mg^{2+}$  ions at high (50 nM) protein concentrations appears to have an opposite effect on binding to undamaged DNA. As seen in Table 2 (A and B), the binding of Spo-UVDE and Bsu-UVDE to the undamaged DNA in the presence of  $Mg^{2+}$  is reduced from 9 % (Spo-UVDE) and 19 % (Bsu-UVDE) to 3 % and 7 % respectively. When both  $Mn^{2+}$  and  $Mg^{2+}$  ions are present, the binding of Spo- and Bsu-UVDE to undamaged DNA is reduced compared to the incubation with  $Mn^{2+}$  alone (Table 2). Taken together, these data indicate that for both Spo-UVDE and Bsu-UVDE  $Mg^{2+}$  prevents their unspecific binding, which might explain the stimulating activity of this cofactor on the activities of these enzymes.

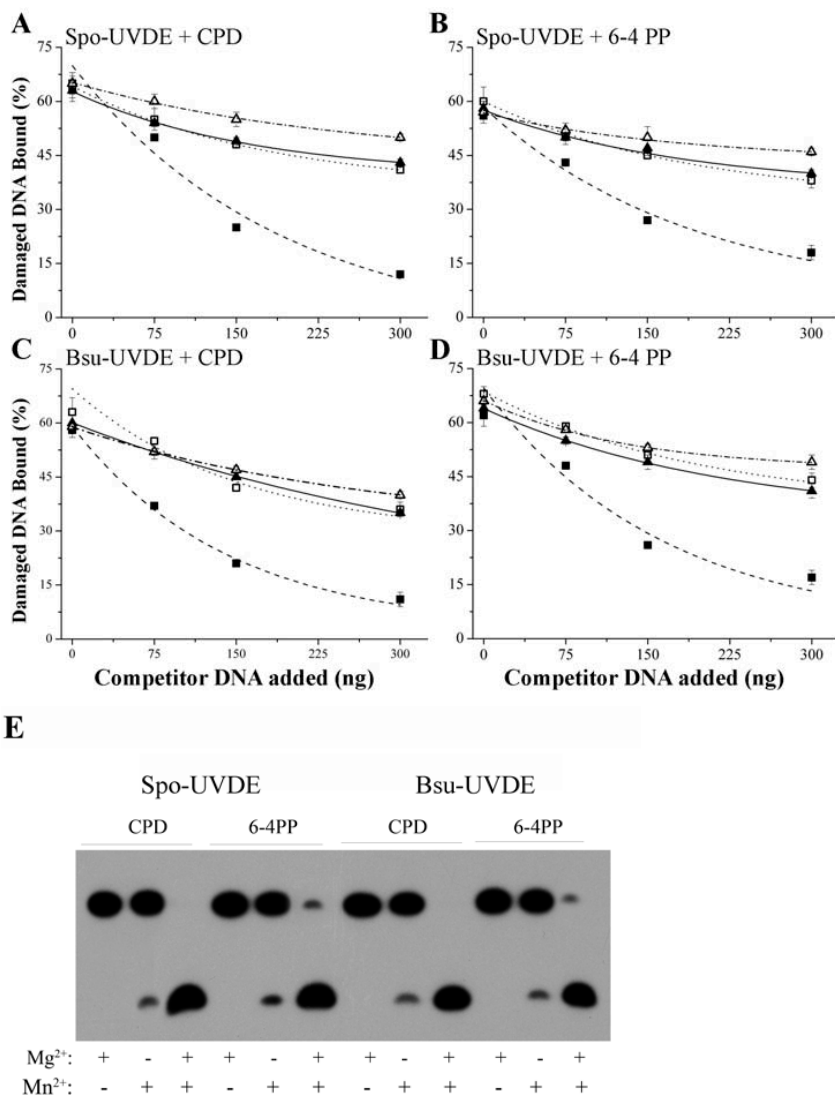
To further investigate the effect of  $Mg^{2+}$  ions on non-specific UVDE binding we performed competition assays. In these assays undamaged competitor DNA (pUC18) was added in increasing amounts to radioactively labeled DNA fragments containing a CPD (Figure 4, A and C) or 6-4PP (Figure 4, B and D). For both Spo-UVDE and Bsu-UVDE in the absence of divalent metal ions the undamaged competitor slightly reduces the amount of damage-specific complexes formed with CPD and the 6-4PP. When 10 mM  $Mg^{2+}$  was included in the reaction the influence of the undamaged DNA is decreased on all tested substrates and the amount of protein-damaged DNA complexes is higher than without metal ions. In the presence of 1 mM  $Mn^{2+}$  the amount of specific complexes drastically decreases indicating a  $Mn^{2+}$  stimulated binding to the competitor DNA. When both  $Mn^{2+}$  and  $Mg^{2+}$  ions were included in the reaction mixes the complex formation of Spo-UVDE and Bsu-UVDE on the CPD and the 6-4PP is restored to the level of damage specific binding in the absence of metal ions. This experiment clearly shows that  $Mn^{2+}$  severely reduces the ability of the UVDE proteins to discriminate between damaged and undamaged DNA and that the addition of  $Mg^{2+}$  ions restores the discriminating properties.

We also tested the UVDE-mediated incision in the presence of non-damaged competitor DNA (Figure 4E). Also in this incision assay the beneficial effect of the ions is clearly visible, since the catalytic activities of both nucleases on the CPD and the 6-4PP are considerably higher with 1 mM  $Mn^{2+}$  and 10 mM  $Mg^{2+}$  than with  $Mn^{2+}$  alone. Under these conditions  $Mg^{2+}$  alone cannot support any nicking activity showing once again that it is not a catalytic cofactor for Bsu-UVDE and Spo-UVDE.

Summarizing we can conclude that the beneficial influence of  $Mg^{2+}$  on the Spo-UVDE and Bsu-UVDE activities is due to destabilizing the proteins complexes on undamaged DNA sites and not by being part of a mixed  $Mn^{2+}/Mg^{2+}$  catalytic site.

### **$Mg^{2+}$ inhibits the Tth-UVDE active site**

When we tested the incision activity of Tth-UVDE in the presence of  $Mn^{2+}$  and  $Mg^{2+}$ , we did not observe a stimulating effect of  $Mg^{2+}$ . Surprisingly on the CPD lesion even an opposite effect was found, since incision was severely reduced in the presence of  $Mn^{2+}$  and  $Mg^{2+}$  compared to  $Mn^{2+}$  alone (Figure 3B, lanes 1 - 4). The inhibitory effect of  $Mg^{2+}$  is restricted to the CPD lesion, since as shown in Figure 3B, the nicking activity of this enzyme on the 6-4PP (lanes 5 - 8) and the abasic site (lanes 9, 10) is not influenced by  $Mg^{2+}$ .



**Figure 4:** Damage specific binding of Spo-UVDE and Bsu-UVDE in the presence of undamaged competitor DNA.

Radioactively labeled DNA (0.5 nM) containing a CPD (A, C) or a 6-4PP (B, D) was incubated with increasing amounts undamaged cold competitor DNA (pUC18) and 2.5 nM Spo-UVDE (A, B) or Bsu-UVDE (C, D). The DNA was incubated for 10 minutes at 30°C in the absence divalent metals (▲), in the presence of 10 mM Mg<sup>2+</sup> (△), in the presence of 1 mM Mn<sup>2+</sup> (■) and in the presence of both 1 mM Mn<sup>2+</sup> and 10 mM Mg<sup>2+</sup> (□).

E. DNA fragments (0.5 nM) containing CPD or 6-4PP were incubated with 300 ng undamaged competitor DNA (pUC18) and 2.5 nM Spo-UVDE or Bsu-UVDE in the presence of 1 mM Mn<sup>2+</sup> and/or 10 mM Mg<sup>2+</sup>, as indicated. The reactions were stopped after 10 minutes at 30°C and the incision products analysed on a 15 % acrylamide gel.

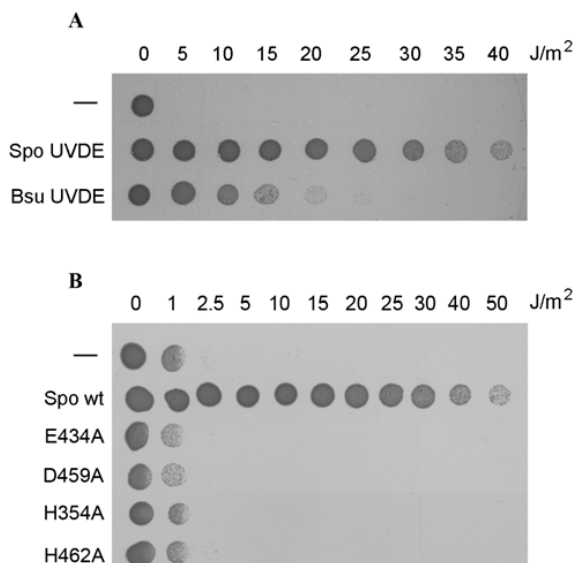
To discriminate if the influence of Mg<sup>2+</sup> on Tth-UVDE is pure catalytical or due to a reduced protein-DNA binding, we performed filter binding tests. As discussed before (Table 2C), in the absence of additional metal ions Tth-UVDE appears to be a weak binder compared to the Spo and Bsu enzymes. Addition of 10 mM Mg<sup>2+</sup> severely decreases the Tth-UVDE complex formation on the CPD from 34 % to 3 %. Mn<sup>2+</sup> can not rescue the disturbed DNA binding, since only 4 % complexes are observed when both Mn<sup>2+</sup> and Mg<sup>2+</sup> are added to the system. Possibly, Mg<sup>2+</sup> binds within the catalytic site thereby competing with Mn<sup>2+</sup>. Since the metal cofactors likely participate in the orientation of the DNA substrate within the protein active site this might result in disturbed DNA binding. Strikingly Mg<sup>2+</sup> can compete with Mn<sup>2+</sup> even at equimolar concentrations (Figure 3B, lane 2).

The Tth-UVDE binding to the 6-4PP and the abasic site is also reduced with Mg<sup>2+</sup> alone (Table 2C). In contrast to the CPD lesion, however, Mn<sup>2+</sup> rescues the 6-4PP and the abasic site binding, since when both Mn<sup>2+</sup> and Mg<sup>2+</sup> ions are present, the binding efficiencies are comparable to those with Mn<sup>2+</sup> alone or in the absence of cofactors (Table 2C).

These data imply that there is a significant difference in the stability of metal ions coordination when Tth-UVDE interacts with CPD on one hand or with 6-4PP and the abasic site on the other. It is very likely that when UVDE is bound to the 6-4PP and the abasic site a stable Mn<sup>2+</sup> coordination prevents Mg<sup>2+</sup> from binding in the enzyme metal coordinating sites. In the UVDE-CPD complex, however, Mg<sup>2+</sup> is capable of removing stable coordination, not only reducing the catalytic efficiency but also the stability of the protein-DNA complexes.

### **The three UVDE proteins differ in their substrate specificities**

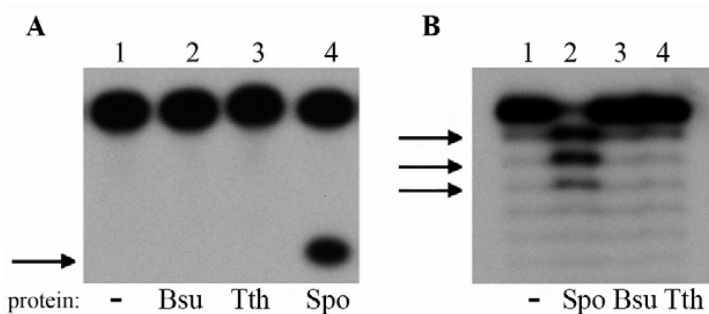
We have previously shown that in *E. coli* the repair of UV lesions by the Spo-UVDE protein is very efficient and equals the repair by the UvrABC Nucleotide Excision Repair proteins (Paspaleva *et al.*, 2009). To check if that is also the case for the bacterial UVDE homologues we compared the UV survival of NER defective *E. coli* strains expressing Spo-UVDE and Bsu-UVDE. Unfortunately, the Tth homologue could not be tested for *in vivo* activity in *E. coli*, since *in vitro* data showed that this protein is not active at 37°C (data not shown). As seen in Figure 5A, the Spo-UVDE gives rise to a high level of UV resistance, while the Bsu enzyme only partially complements the UV sensitivity of the NER deficient strain. The observed effect is surprising, since both proteins are expressed to the same level (not shown) and they exhibited comparable activity on the 6-4PP and the CPD lesions *in vitro* (Figure 3C). One possibility is that the *in vitro* incubation conditions (pH 6.5) differ from the *in vivo* situation and therefore we tested the activity of both UVDEs at a more physiological pH (7.5). These tests, however, also did not reveal a clear difference between the two enzymes (data not shown). Another explanation for the observed *in vivo* differences might be that the substrate specificity of Spo-UVDE might be broader than that of Bsu-UVDE, including minor UV lesions (other than CPD and 6-4PP). To compare the substrate specificities of Spo-UVDE and Bsu-UVDE we analysed the activities of these enzymes on two other types of DNA modifications, a thymine glycol and DNA containing a single strand nick. These substrates have previously been reported to be incised by the



**Figure 5:** UV survival of *E. coli* strains expressing Spo-UVDE or Bsu-UVDE proteins.

NER deficient ( $\Delta uvrA$ ,  $\Delta uvrB$ ) *E. coli* cells containing a plasmid expressing the indicated wild type (A) or mutant (B) UVDE proteins were irradiated with different doses of UV light. The empty pUC plasmid was used as a negative control (upper rows).

Spo-UVDE (Paspaleva *et al.*, 2009). As seen in Figure 6A, Spo-UVDE shows a moderate activity on the thymine glycol containing fragment (lane 4), but Bsu-UVDE completely fails to incise the DNA containing this damage (lane 3). Furthermore, the Bsu protein also doesn't cut the DNA containing ss-nick (Figure 6B, lane 3), which is a good substrate for the Spo-UVDE (Figure 6B, lane 2).



**Figure 6:** UVDE-mediated incision of DNA substrates containing a thymine glycol or a single strand nick.

A. Tth-, Spo- and Bsu-UVDE enzymes (50 nM) were incubated for 15 minutes with 0.2 nM DNA fragment containing a thymine glycol.

B. DNA containing a ss nick was incubated with 10 nM of the same UVDE proteins. The arrows indicate the position of the incision products.

Despite the lack of incision activity, the binding of the Bsu homologue to the thymine glycol is not significantly different from the complex formation on the CPD and the 6-4PP (Table 3). The complex formation of Bsu-UVDE on single strand nick is reduced, but still significantly higher than on the undamaged DNA. Taken together these findings suggest that Bsu-UVDE is still able to recognize the thymine glycol and the DNA nick as damages, but fails to properly position them in its active site. The same could be true for other types of lesions like other UV induced DNA damages, which might explain the reduced *in vivo* activity of Bsu-UVDE.

**Table 3:** Binding efficiencies of UVDE enzymes on different DNA lesions.

DNA damage:	Spo	Bsu	Tth
No	3 ± 1	3 ± 2	3 ± 1
CPD	55 ± 3	55 ± 2	34 ± 2
6-4PP	57 ± 1	56 ± 3	19 ± 1
Abasic site	53 ± 3	60 ± 1	20 ± 2
Thymine glycol	58 ± 1	50 ± 2	4 ± 1
DNA nick	52 ± 1	39 ± 2	3 ± 3

DNA with the indicated damage was incubated with 50 nM Tth-UVDE, 5 nM Spo-UVDE or 5 nM Bsu-UVDE for 10 minutes in the absence of divalent metal ions.

We also tested the Tth-UVDE on the thymine glycol and DNA containing a single strand nick. As seen in Figure 6 A, B (lane 4), it not only completely failed to incise the two substrates but it can no longer discriminate this damage from undamaged DNA (Table 3). Since the Tth protein outlines itself as a weak binder even on substrates introducing major changes in the DNA structure such as the 6-4PP, the structural changes induced by a nick or a thymine glycol most likely are not sufficient to promote proper DNA binding.

In conclusion, Spo-UVDE distinguishes itself from its bacterial homologues with a much broader substrate specificity. Bsu-UVDE is more restricted, although in addition to the CPD and 6-4PP, it can also incise an abasic site. The Tth enzyme, however, can only efficiently repair CPD and 6-4PP lesions.

### Dynamics of Spo and Bsu UVDE pre-incision complexes

Previously we have shown by using 2-aminopurine (2-AP) fluorescence studies that the Spo-UVDE causes significant destacking of the two bases opposite a CPD and a 6-4PP (Paspaleva *et al.*, 2009). Similar conformational changes were observed with the base facing the abasic site. Furthermore, quenching experiments with acrylamide revealed that these bases in the non-damaged strand do not become solvent exposed and are in a contact with side chains of the protein.

To see if similar conformational changes occur upon Bsu-UVDE binding to these damages we compared the 2-AP fluorescence in the presence of this enzyme to that induced by Spo-UVDE. For this purpose we used DNA fragments containing 2-AP incorporated at two positions

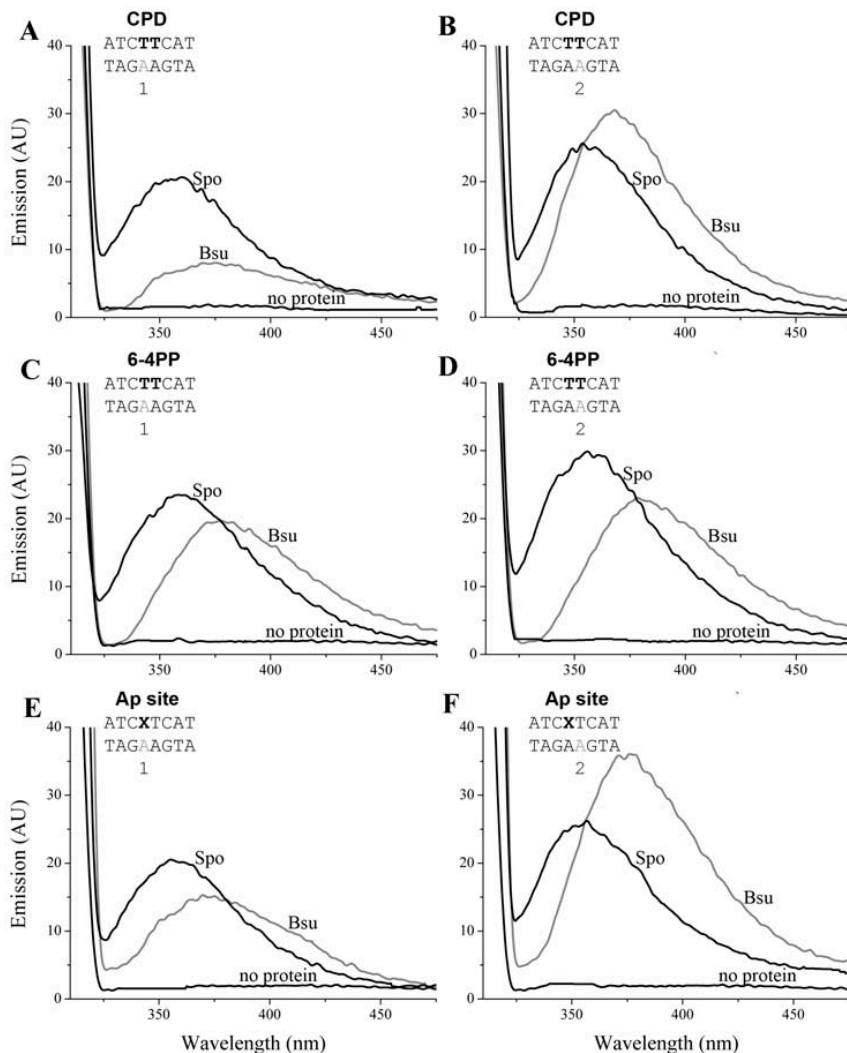
opposite a CPD, 6-4PP or an abasic site. The presence of the 2-AP did not change the incision efficiencies of either Bsu-UVDE or Spo-UVDE (data not shown). The binding properties of both nucleases are also not altered by the presence of the 2-AP (Table 4).

**Table 4:** *S. pombe* and *B. subtilis* UVDE binding on 2-AP containing DNA substrates.

Damage	2-AP position	% binding Spo	% binding Bsu
CPD	no	61 ± 1	62 ± 1
CPD	1	60 ± 1	62 ± 1
CPD	2	59 ± 1	57 ± 3
6-4PP	no	59 ± 3	63 ± 1
6-4PP	1	61 ± 4	61 ± 3
6-4PP	2	65 ± 1	64 ± 2
Abasic site	no	62 ± 1	62 ± 2
Abasic site	1	61 ± 4	61 ± 1
Abasic site	2	60 ± 1	62 ± 3
Thymine glycol	no	60 ± 1	58 ± 2
Thymine glycol	1	60 ± 2	57 ± 2
Thymine glycol	2	61 ± 1	58 ± 1
Nick	no	57 ± 1	54 ± 2
Nick	1	57 ± 2	55 ± 3
Nick	2	59 ± 1	58 ± 1

The incubations were done using the same conditions as used for the fluorescence assays without metal cofactors. The binding is expressed as the percentage of total input DNA retained on a filter.

In the absence of protein the CPD, the 6-4PP and the abasic site lesion containing fragments do not exhibit significant fluorescence (Figure 7, A - F), showing that these damages do not disturb the base stacking in the opposing strand. As previously shown (Paspaleva *et al.*, 2009), the 2-AP signal in the presence of Spo-UVDE is significantly increased at both positions. In comparison, the Bsu-UVDE also causes notable change in the 2-AP fluorescence, suggesting that like Spo-UVDE this homologue is also able to cause destacking of the bases opposite the DNA damage. There are, however, multiple differences between Spo-UVDE and Bsu-UVDE. The levels of destacking differ depending on the damage. For instance upon incubation of the CPD and the AP site containing DNA with Bsu-UVDE the fluorescence of the 2-AP at position 1 is lower than the one caused by the Spo homologue (Figure 7, A and E), whereas the signal at position 2 is notably higher (Figure 7, B and F). Both proteins, however, cause a more pronounced effect on position 2. On the 6-4PP containing fragment Bsu-UVDE has a higher impact on the 2-AP at position 1, compared to the CPD and the AP site containing fragments, but the signal is still lower compared to Spo-UVDE (Figure 7C). Also the signal at position 2 is now lower than that of Spo-UVDE (Figure 7D). Another striking difference between the Spo and the Bsu enzymes is that for both 2-AP residues the emission maximum in the Bsu-UVDE and Spo-UVDE complexes are shifted with respect to each other by ~ 50 nm.



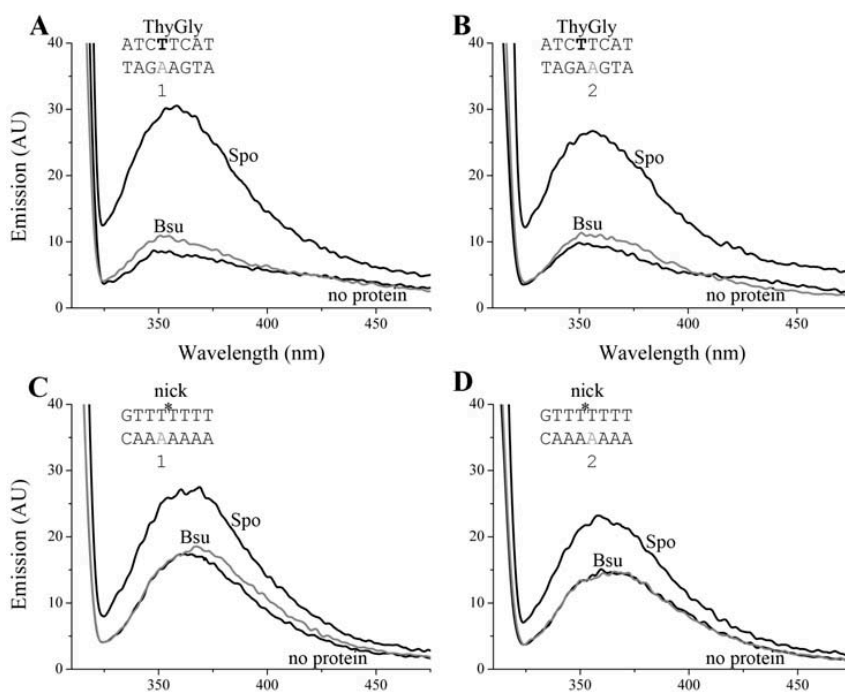
**Figure 7:** Fluorescence emission spectra of UVDE-DNA complexes.

The DNA fragments (0.5  $\mu$ M) with a CPD (A and B), 6-4PP (C and D) or an abasic site (E and F) were incubated with or without 2.5  $\mu$ M Spo-UVDE and Bsu-UVDE for 10 minutes at 30°C. Panels A, C and E show the 2-AP signal when the fluorescent base is positioned directly opposite the 5'T of the CPD and the 6-4PP or opposite the abasic site, while in panels B, D and F the 2-AP is incorporated opposite the 3'T of the CPD, 6-4PP and the 3' neighbor of the abasic site.

When we added the quenching agent acrylamide in our incubation mixes we couldn't detect any change in the 2-AP fluorescence in either of the Spo-UVDE or Bsu-UVDE complexes (data not shown). Apparently in both enzymes the bases in the non-damaged strand are shielded from the solution by protein side chains. The way that this is achieved, however, might be different

for the two proteins as illustrated by the shift in emission maximum, which is indicative of a different polarity in the environment of the 2-AP residues (Ward *et al.*, 1969).

Next we compared the activity of both enzymes on DNA fragments where the 2-AP was incorporated opposite a thymine glycol and ssDNA nick. As shown before (Paspaleva *et al.*, 2009), although the CPD, the 6-4PP and the abasic site do not cause destacking in the undamaged strand by themselves this is not the case for a ss nick or a thymine glycol, where in an absence of protein a fluorescent signal of 15 AU and 10 AU, respectively is observed (Figure 8). The Spo-UVDE causes a significant destabilization in the undamaged DNA strand of a thymine glycol containing DNA (Figure 8, A and B) and again addition of acrylamide did not reduce the fluorescence signal (Paspaleva *et al.*, 2009). In contrast the Bsu-UVDE doesn't cause any detectable destacking in the undamaged strand of the thymine glycol containing fragment (Figure 8, A and B). This is not due to a lack of protein-DNA complexes, since filter binding studies revealed that at the conditions used 58 % of the DNA is bound by the Bsu-UVDE (Table 4).



**Figure 8:** Fluorescence emission spectra of UVDE-DNA complexes.

DNA (0.5  $\mu$ M) containing a thymine glycol (A and B) or ss nick (C and D) was incubated with 2.5  $\mu$ M Spo-UVDE and 4  $\mu$ M Bsu-UVDE for 10 minutes at 30°C. In panels A, C the 2-AP is placed opposite the base flanking the nick on the 5' side or directly opposite the thymine glycol. In panels B, D the fluorescent base faces the nucleotide on the 3' side of the nick or the 3' neighbor of the thymine glycol.

A similar result was obtained with the two 2-AP residues opposite the two T's flanking a nick (Figure 8, C and D). Whereas Spo-UVDE does cause destacking of these bases, Bsu-UVDE fails to induce any change in fluorescence despite the 55 % binding efficiency (Table 4). Apparently Bsu-UVDE does recognize a DNA fragment containing a thymine glycol or a nick, but is not able to induce conformational changes in the opposing strand, which might explain why this protein fails to incise these DNA damages.

We also performed fluorescence studies with the Tth-UVDE, however no change of the 2-AP signal was detected on any of the substrates. The amount of complexes detected under the fluorescence conditions (~ 15 %) is probably too low.

### Active site organization of Spo UVDE

Based on sequence alignment of Spo-UVDE, Bsu-UVDE and the known metal coordinating residues in Tth-UVDE, it is likely that in Spo-UVDE Mn1 is coordinated by Glu434, Glu535, His498, Asp459 (Figure 1). Mn2 is oriented with the help of Glu434, His354 and His401 and according to the alignment Mn3 must be in a contact with Met511 and His462. Methionine, however, is very unlikely to act as a metal coordinator and possibly in Spo-UVDE this role is performed by His516. All metal coordinators, except the one replaced by methionine in the Spo-UVDE sequence, are highly conserved (Figure 1). The importance of Mn1 for Spo-UVDE function was previously demonstrated, since alanine substitutions of Glu434 and Asp459 abrogated the catalytic activity of the enzyme (Paspaleva *et al.*, 2007). To test the importance of the other metals we made additional substitutions of His354 and His462. As expected the resulting mutant proteins are disturbed in incising DNA containing an abasic site (Table 5) with H462A showing a 50 % reduction and H354A completely failing to incise the abasic site lesion. On the CPD and the 6-4PP, however, the activity of the mutants is comparable to the wild type protein (Table 5). Once again this shows that the metal coordination is influenced by the type of lesion. Apparently in the mutant proteins the three metal ions can still be properly positioned at the UV-lesions but not on the abasic site.

Surprisingly all Spo-UVDE active site mutants, including H354A and H462A appeared completely repair deficient *in vivo* (Figure 5B). One possible explanation for the discrepancy between the *in vivo* and the *in vitro* data might be that the conditions used in our incision assays

**Table 5:** Effect of the pH on the incision activity of wild type and mutant UVDE enzymes.

UVDE	CPD		6-4PP		Abasic site pH 6.5
	pH 6.5	pH 7.5	pH 6.5	pH 7.5	
wild type	97 ± 2	97 ± 1	95 ± 1	95 ± 2	95 ± 1
H354A	97 ± 3	20 ± 1	95 ± 2	20 ± 3	<1
E434A	<1	<1	<1	<1	<1
D459A	<1	<1	<1	<1	<1
H462A	97 ± 1	10 ± 3	95 ± 3	10 ± 1	50 ± 4

Incubation was done with 5 nM UVDE enzymes and 0.2 nM DNA in the presence of 1 mM Mn<sup>2+</sup> and 10 mM Mg<sup>2+</sup> at pH 6.5 and pH 7.5. The numbers represent the incision percentages.

differ from the *in vivo* situation. The cellular pH of *E. coli* is around 7.5 (Padan *et al.*, 1981). In literature the optimal pH for the Spo-UVDE was reported to be 6.5 (Kaur *et al.*, 1998) and all incision and DNA binding assays have therefore been performed in a buffer of this lower pH. To test if indeed the pH might be an explanation for the observed differences we tested the incision efficiency of Spo-UVDE at pH 7.5 and pH 6.5. As shown before (Paspaleva *et al.*, 2009) the higher pH does not influence the incision activity of the wild type Spo enzyme on any of the tested lesions (Table 5). However, when we tested the activity of Spo active site mutants at pH 7.5, the incision activities of both H354A and H462A are significantly reduced (Table 5). The damage-specific binding of the mutants, however, is not altered by the pH (Table 6). Apparently the metal coordination in the UVDE active site is strongly influenced by the pH, most likely by influencing the position of the amino acid residues, which participate in the metal ions coordination.

**Table 6:** Effect of the pH on the DNA binding properties of wild type and mutant Spo-UVDE enzymes.

UVDE	Undam. DNA *		CPD		6-4PP		Abasic site	
	pH 6.5	pH 7.5	pH 6.5	pH 7.5	pH 6.5	pH 7.5	pH 6.5	pH 7.5
wild type	9 ± 1	6 ± 3	55 ± 3	52 ± 2	57 ± 1	57 ± 4	56 ± 1	59 ± 2
H354A	11 ± 2	9 ± 3	52 ± 4	53 ± 1	60 ± 2	57 ± 3	60 ± 2	58 ± 2
E434A	7 ± 3	5 ± 2	53 ± 1	49 ± 1	59 ± 2	56 ± 3	61 ± 1	57 ± 2
D459A	10 ± 2	7 ± 1	58 ± 1	54 ± 4	61 ± 3	58 ± 1	59 ± 1	62 ± 2
H462A	7 ± 1	6 ± 1	54 ± 1	51 ± 3	61 ± 2	61 ± 2	58 ± 2	62 ± 1

\*Incubation on the undamaged DNA was done with 50 nM Spo UVDE and 4 nM DNA.

The incubation on the CPD, 6-4PP and the abasic site lesion was performed with 5 nM UVDE and 4 nM DNA fragments. All incubations were done in the absence of additional metals. The binding is expressed as the percentage of total input DNA.

## DISCUSSION

The recently solved UVDE structure from *T. thermophilus* (Paspaleva *et al.*, 2007) outlined the presence of three divalent metals and crystallisation trials in the presence of 1 mM Mn<sup>2+</sup> showed that the three metal binding sites have the potential to be all occupied by Mn<sup>2+</sup>, forming a compact cluster (Meulenbroek *et al.*, 2009). However, it was not clarified if Mn<sup>2+</sup> is indeed the enzyme's optimal catalytic cofactor. In our study we not only give direct evidence that Mn<sup>2+</sup> is the actual UVDE cofactor, but we also show which other metals can substitute the Mn<sup>2+</sup> ions. In our tests Co<sup>2+</sup> is revealed to be one of the most effective substitutes for Mn<sup>2+</sup>, not only for the Spo protein, but also for Tth-UVDE and to a lesser extend for Bsu-UVDE. Notably Cu<sup>2+</sup> and Zn<sup>2+</sup>, which are highly similar in ionic radius and exhibit the same charge as Co<sup>2+</sup> can support only a very low enzymatic activity, thus the ionic radius, although important, is not the only decisive factor for the metal function.

Spo-UVDE distinguishes itself from Bsu- and Tth-UVDE by being able to use  $Ca^{2+}$  as a metal cofactor. The large ionic radius of  $Ca^{2+}$  makes it highly unlikely that a cluster of three  $Ca^{2+}$  ions is formed in the catalytic site of UVDE. Possibly  $Ca^{2+}$  can be part of mixed  $Mn^{2+}/Ca^{2+}$  active site, making use of one or more intrinsically bound  $Mn^{2+}$  ions. An active complex, however, could only be formed on the 6-4PP indicating that the DNA itself has a major role in the coordination of the metal ions (see also below).

Both Bsu and Tth proteins exhibit more stringent divalent metal requirements compared to the Spo orthologue. From the three nucleases Tth-UVDE has the most restricted metal requirement as incision is supported only in the presence of  $Mn^{2+}$  or  $Co^{2+}$ . However, bearing in mind that binding of divalent metals is expected to be temperature dependent, the incubations of the thermophilic protein at  $55^{\circ}C$  might negatively influence stable coordination of other cofactors.

An intriguing observation is that most of the tested divalent metal ions have been found to support different levels of UVDE activity, depending on the type of damage in the DNA substrate. This indicates that the DNA itself assists in the coordination of the metal ions, implying that at least part of these metal ions bind into the active site after association of the protein with the DNA. The crystal structure of the Tth-UVDE protein revealed the presence of a phosphate ion coordinated by all three  $Mn^{2+}$  ions (Paspaleva *et al.*, 2007). Most likely this phosphate ion represents a phosphate of the DNA backbone in the protein-DNA complex. Positioning of this DNA-phosphate is likely to be dependent on the type of DNA lesion explaining why metal coordination is dependent on the DNA substrate. This model is supported by the properties of the Spo-UVDE active site mutant H354A (coordinator of  $Mn^{2+}$ ), which is completely deficient in nicking the abasic site but shows wild type level of incision on the CPD and the 6-4PP, again showing that metal coordination is stabilised via the DNA.

In our studies we were also able to distinguish between metal ions needed to occupy the active site and metal ions affecting the DNA binding.  $Mg^{2+}$  although used by some authors in UVDE incision assays (Kanno *et al.*, 1999), can support incision activity only in the case of the Spo homologue and even then a higher concentration of  $Mg^{2+}$  ions (10 mM) is needed. One likely explanation is the reported high charge density of  $Mg^{2+}$ , which poses a strong electrostatic barrier for formation of a three ions  $Mg^{2+}$  cluster (Cowan, 2002). The fact that Spo-UVDE can still use this metal ion for catalysis suggests that a mixed  $Mg^{2+}/Mn^{2+}$  metal binding site is possible for this nuclease. This in combination with the ability of Spo-UVDE to perform incision in the presence of numerous other divalent metals underlines a high flexibility of its active site.

In our DNA binding experiments we showed that  $Mg^{2+}$  has a role in preventing the binding of Spo- and Bsu-UVDE to undamaged DNA, thereby enhancing its damage-specific incision. Therefore for both nucleases the combination of  $Mn^{2+}$  and  $Mg^{2+}$  provides the optimal reaction condition:  $Mn^{2+}$  acting as a pure catalytic cofactor and  $Mg^{2+}$  as a DNA binding modifier. A reducing effect of  $Mg^{2+}$  on non-specific DNA binding has also been described for other enzymes, like the B-ZIP transcription factor (Moll *et al.*, 2002) and the restriction nuclease *EcoRV* (Jeltsch *et al.*, 1997).  $Mg^{2+}$  has been shown to bind to the phosphates of the DNA backbone (Cowan,

1998; Hackl *et al.*, 2005) and needs to be released upon protein binding. The interaction of UVDE at non-specific sites is expected to be mainly via these phosphates, hence interference by the  $Mg^{2+}$  ions. Damage specific binding, as expected for a non-sequence specific protein, will also involve phosphate contacts. In addition, however, the complex will be stabilised by other factors like DNA kinking and specific interactions with the damaged bases in the active site (see also below), compensating for the inhibitory effect of  $Mg^{2+}$ .

$Mn^{2+}$  appeared to have an opposite effect on DNA binding as it was seen to stimulate binding to undamaged DNA. Although  $Mn^{2+}$  also binds to DNA it was shown to have a higher affinity for the DNA bases causing destabilisation and bending of the DNA (Polyanichko *et al.*, 2004; Hackl *et al.*, 2005). Since, as described before (Paspaleva *et al.*, 2009) UVDE binds to sites, which allow DNA bending,  $Mn^{2+}$  might promote unspecific binding of the protein by altering the DNA flexibility.

For Tth-UVDE the additional presence of  $Mg^{2+}$  resulted in a severe reduction of incision on the CPD, whereas the activity on the 6-4PP was not affected. It is very unlikely that the negative effect of  $Mg^{2+}$  is due to an influence on specific protein-DNA contacts (i.e. competing for phosphate binding), since in the absence of cofactors the protein-DNA complex on the CPD is even more stable than on the 6-4PP. More likely  $Mg^{2+}$  becomes part of a mixed  $Mg^{2+}/Mn^{2+}$  active site leading to a catalytically inactive protein. Since, as discussed above, metal coordination is highly dependent on the type of substrate the disturbing coordination of  $Mg^{2+}$  does not take place on the 6-4PP. Strikingly the hampering effect of  $Mg^{2+}$  on CPD incision was already apparent at equimolar concentrations of 1 mM, making it questionable whether *in vivo* the Tth-UVDE would be capable of repairing CPD at all. Although the concentration of free  $Mg^{2+}$  in the *T. thermophilus* cells is rather low: 1.53 mM (Kondo *et al.*, 2007) or according to other authors (Wakamatsu *et al.*, 2007) even lower (0.53 mM), it is not very likely that it will be greatly exceeded by the concentration of  $Mn^{2+}$  (which until now has not been determined). In that regard it is possible that *in vivo* the UVDE enzyme of *T. thermophilus* will only repair 6-4PP lesions.

The three UVDE homologues show very different *in vitro* substrate specificity. Spo-UVDE is not only active on CPD, 6-4PP and abasic sites, but also on thymine glycol and DNA containing a nick. We have, however, previously shown (Paspaleva *et al.*, 2009) that incision of the DNA containing a nick or abasic site is highly dependent on the presence of flanking pyrimidines. It was therefore proposed that in order to position the scissile phosphodiester bond into the active site, the damaged nucleotides have to be rotated into a pocket of the protein, which is located at the bottom of a proposed DNA binding groove (Paspaleva *et al.*, 2007). Purine residues will be excluded from this pocket making the protein selective for (distorted) pyrimidines or an abasic site together with a pyrimidine. Most likely the complex will be further stabilised by contacts of the rotated pyrimidine(s) with side chains of the protein pocket. The substrate specificity of the Bsu homologue appears to be more limited. Although the DNA containing a CPD, 6-4PP or abasic site are efficiently incised no activity on DNA with a thymine glycol or single strand nick

could be detected. A clue for what might cause the difference between the two proteins came from studying conformational changes in the opposing strand using the fluorescent residue 2-AP. Previously, it was shown with Spo-UVDE that Tyr358 causes destacking of bases opposing the lesion. A model was proposed in which Tyr358 will wedge between the two residues opposite the damage thereby disturbing the base stack and stabilizing the DNA kink (Paspaleva *at al.*, 2009). With Bsu-UVDE a similar destacking of bases opposite the CPD, 6-4PP or abasic site was observed but the protein appeared incapable of inducing a conformational change of bases opposing a nick or thymine glycol. Tyr358 is part of a GQY loop, which in the crystal structure of Tth-UVDE is located at the bottom of a proposed DNA binding groove. Although this loop is conserved in the Spo and Tth proteins the sequence of the corresponding region in Bsu-UVDE is NQF (Figure 1). Although the corresponding phenylalanine in Bsu-UVDE can fulfil a similar wedging function as Tyr358, the positioning of this residue might be different as a result of the presence of an asparagine instead of a glycine in the NQF loop. That indeed on the CPD, 6-4PP and abasic site the Phe residue of Bsu-UVDE is positioned differently was not only evident from a difference in the level of the fluorescence signals induced by the Spo- and Bsu proteins. More strikingly the emission maximum of the 2-AP residues was shifted by about 50 nm indicating that the bases in the non-damaged strand are placed in a different amino acid context. Possibly this different NQF loop could also be the reason that Bsu-UVDE is incapable of causing destacking of the bases opposing a thymine glycol or single-strand nick. As a consequence also the damage itself is not properly positioned into the active site of the protein and subsequent incision cannot take place.

The Tth-UVDE shows the most restricted substrate specificity, only efficiently incising the CPD and 6-4PP. The protein-DNA complexes formed by Tth-UVDE appeared much more instable so we could not test whether also for this protein the substrate specificity is related to the capacity to induce destacking of bases in the non-damaged strand. Because the high temperature used for Tth-UVDE will inflict more flexibility in the DNA, this protein might be more dependent on direct interactions with the damaged nucleotides for proper positioning of the scissile phosphodiester bond. The absence of an interacting base in the abasic site substrate could then explain its lower incision efficiency on this DNA.

## REFERENCES

- Bowman KK, Sidik K, Smith CA, Taylor JS, Doetsch PW, Freyer GA, A new ATP-independent DNA endonuclease from *Schizosaccharomyces pombe* that recognizes cyclobutane pyrimidine dimers and 6-4 photoproducts, *Nucleic Acids Res.* **22** (1994), pp. 3026-3032
- Cowan JA, Structural and catalytic chemistry of magnesium-dependent enzymes, *Biometals* **15** (2002), pp. 225-235
- Doetsch PW, Beljanski V and Song B, The ultraviolet damage endonuclease (UVDE) protein and alternative excision repair: a highly diverse system for damage recognition and processing. In: V. Beljanski and B. Song, Editors, *DNA Damage Recognition*, Taylor & Francis Press, New York (2006), pp. 211-223
- Evans DM, Moseley BE, Identification and initial characterisation of a pyrimidine dimer UV endonuclease (UV endonuclease beta) from *Deinococcus radiodurans*; a DNA-repair enzyme that requires manganese ions, *Mutat Res.* **145** (1985), pp. 119-128
- Goosen N, Moolenaar GF, Repair of UV damage in bacteria, *DNA Repair* **7** (2008), pp. 353-379
- Hackl EV, Kornilova SV, Blagoi YP, DNA structural transitions induced by divalent metal ions in aqueous solutions, *Int J Biol Macromol.* **35** (2005), pp. 175-191
- Henne A, Bruggemann H, Raasch C, Wiezer A, Hartsch T, Liesegang H, Johann A, Lienard T, Gohl O and Martinez-Arias R, *et al.*, The genome sequence of the extreme thermophile *Thermus thermophilus*, *Nat. Biotechnol.* **22** (2004), pp. 547-553
- Hosfield DJ, Guan Y, Haas BJ, Cunningham RP and Tainer JA, Structure of the DNA repair enzyme Endo IV and its DNA complex: double-nucleotide flipping at abasic sites and three-metal-ion catalysis, *Cell* **98** (1999), pp. 397-408
- Iwai S, Chemical synthesis of oligonucleotides containing damaged bases for biological studies, *Nucleosides Nucleotides Nucleic Acids* **25** (2006), pp. 561-582
- Jeltsch A, Maschke H, Selent U, Wenz C, Köhler E, Connolly BA, Thorogood H, Pingoud A, DNA binding specificity of the *EcoRV* restriction endonuclease is increased by  $Mg^{2+}$  binding to a metal ion binding site distinct from the catalytic center of the enzyme. *Biochemistry* **34** (1995), pp. 6239-6246
- Kanno S, Iwai S, Takao M, Yasui A, Repair of apurinic/apyrimidinic sites by UV damage endonuclease; a repair protein for UV and oxidative damage, *Nucleic Acids Res.* **27** (1999), pp. 3096-3103
- Kaur B, Fraser JL, Freyer GA, Davey S, Doetsch PW, A Uve1p-mediated mismatch repair pathway in *Schizosaccharomyces pombe*, *Mol Cell Biol.* **19** (1999), pp. 4703-4710
- Kondo N, Nishikubo T, Wakamatsu T, Ishikawa H, Nakagawa N, Kuramitsu S, Masui R, Insights into different dependence of dNTP triphosphohydrolase on metal ion species from intracellular ion concentrations in *Thermus thermophilus*, *Extremophiles.* **12** (2008), pp. 217-223
- Kunst F, Ogasawara N, Moszer I, Albertini AM, Alloni G, Azevedo V, Bertero MG, Bessieres P, Bolotin A and Borchert S, *et al.*, The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*, *Nature* **390** (1997), pp. 249-256
- Kutyavin IV, Milesi D, Belousov Y, Podyminogin M, Vorobiev A, Gorn V, Lukhtanov EA, Vermeulen NM, Mahoney W, A novel endonuclease IV post-PCR genotyping system, *Nucleic Acids Res.* **34** (2006), pp. 28-134
- Levin JD, Johnson AW, Demple B, Homogeneous *Escherichia coli* endonuclease IV. Characterization of an enzyme that recognizes oxidative damage in DNA, *J Biol. Chem.* **263** (1988), pp. 8066-8071
- Levin JD, Shapiro R, Demple B, Metalloenzymes in DNA repair. *Escherichia coli* endonuclease IV and *Saccharomyces cerevisiae* Apn1, *J Biol. Chem.* **266** (1991), pp. 22893-22898
- Meulenbroek EM, Paspaleva K, Thomassen E, Abrahams JP, Goosen N, Pannu NS, Involvement of a carboxylated lysine in UV damage endonuclease, *Protein Science*, (2009), in press

- Moll JR, Acharya A, Gal J, Mir AA, Vinson C, Magnesium is required for specific DNA binding of the CREB B-ZIP domain, *Nucleic Acids Res.* **30** (2002), 1240-1246
- Moolenaar GF, Visse R, Ortiz-Buysse M, Goosen N, van de Putte P, Helicase motifs V and VI of the *Escherichia coli* UvrB protein of the UvrABC endonuclease are essential for the formation of the preincision complex, *J Mol Biol.* **240** (1994), pp. 294-307
- Padan I, Zilberstein D, Schuldiner S, pH homeostasis in bacteria, *Biochim. Biophys. Acta* **650** (1981), pp. 151-166
- Paspaleva K, Moolenaar GF, Goosen N, Damage recognition by UV damage endonuclease from *Schizosaccharomyces pombe*, *DNA Repair* (2009), in press
- Paspaleva K, Thomassen E, Pannu NS, Iwai S, Moolenaar GF, Goosen N, Abrahams JP, Crystal structure of the DNA repair enzyme ultraviolet damage endonuclease, *Structure* **15** (2007), pp. 1316-1324
- Polyanichko AM, Andrushchenko VV, Chikhirzhina EV, Vorob'ev VI, Wieser H, The effect of manganese(II) on DNA structure: electronic and vibrational circular dichroism studies, *Nucleic Acids Res.* **32** (2004), pp. 989-996
- Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW, Use of T7 RNA polymerase to direct expression of cloned genes, *Methods Enzymol.* **185** (1990), pp. 60-89
- Takao M, Yonemasu R, Yamamoto K and Yasui A, Characterization of a UV endonuclease gene from the fission yeast *Schizosaccharomyces pombe* and its bacterial homolog, *Nucleic Acids Res.* **24** (1996), pp. 1267-1271
- Verhoeven EE, van Kesteren M, Turner JJ, van der Marel GA, van Boom JH, Moolenaar GF and Goosen N, The C-terminal region of *Escherichia coli* UvrC contributes to the flexibility of the UvrABC nucleotide excision repair system, *Nucleic Acids Res.* **30** (2002), pp. 2492-2500
- Wakamatsu T, Ishikawa H, Nakagawa N, Kuramitsu S, Masui R, Intracellular ion concentrations in *Thermus thermophilus*, *Extremophiles* **15** (2008), pp. 317-323
- Ward DC, Reich E, Streyer L, Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside and their derivatives, *J Biol Chem.* **244** (1969), pp. 1228-1237



## SUMMARY AND DISCUSSION

UVDE is a repair enzyme discovered for the first time in the fission yeast *Schizosaccharomyces pombe*. The initial biochemical characterization of this enzyme showed that its substrate specificity includes not only UV lesions, but also abasic sites and some nucleotide mismatches. The mechanism, however, of UVDE damage recognition and DNA cleavage was not clarified. *S. pombe* UVDE was seen to require  $Mn^{2+}$  and  $Mg^{2+}$  for its function, but the utilization of divalent metals in its cleavage reaction was empirical and without a clear vision of the nature, the number and the exact role of the metal cofactors.

UVDE homologues were found in many fungal species and in a number of bacteria such as *Bacillus subtilis* and *Thermus thermophilus*. Although the *S. pombe* UVDE was partially characterized, no information was available for the substrate specificity or the cofactor requirements its bacterial homologues.

The first glimpse at UVDE at molecular level is given by the *T. thermophilus* UVDE structure (Chapter 2). The enzyme is seen to be a single domain TIM-barrel with a wide 29 Å groove and three metals coordinating site buried deeply at bottom of the groove. Except for H224, all the residues involved in the coordination of the metal ions are fully conserved in all known UVDE homologues. Point mutations both in *T. thermophilus* UVDE (Chapter 2) and in *S. pombe* UVDE (Chapter 5) reveal the crucial importance of the divalent metals for the UVDE function, since alanine substitutions of most of the metal coordinating amino acids abrogate activity. The UVDE divalent metals superimpose on the three  $Zn^{2+}$  ions of Endo IV shedding some light on how UVDE might use them to incise DNA. A model is proposed (Chapter 2 and 5) where UVDE uses the metal ion at position 1 and probably 2 to create a hydroxide nucleophile, which attacks the phosphodiester backbone of DNA creating an unstable intermediate. The divalent metal at position 3 is likely to participate mainly in the stabilization of the intermediate step.

Although the fluorescent scan was unable to determine the exact nature of the found metals, biochemical tests suggested that they might be  $Mn^{2+}$  (Chapter 2). That  $Mn^{2+}$  is the actual UVDE cofactor is later confirmed by in detail metal substitution study (Chapter 5). Although  $Mn^{2+}$  is not the only divalent metal that UVDE can use, utilization of other metal ions is less efficient and highly dependent on the type of DNA damage. Furthermore, divalent metals are seen to have an opposite effect on different UVDE enzymes.  $Mg^{2+}$ , which positively influences the activity of *S. pombe* and *B. subtilis* UVDE by preventing their unspecific binding, completely abrogates the *T. thermophilus* incision of the CPD lesion. This unexpected effect of  $Mg^{2+}$  is clearly linked to a reduction in the stability of the protein-CPD complexes. On the 6-4PP, however, the negative influence of  $Mg^{2+}$  is neutralized by  $Mn^{2+}$ , since when both  $Mn^{2+}$  and  $Mg^{2+}$  ions are present the binding efficiency of the *T. thermophilus* UVDE is restored. These data imply that there is significant difference in the cofactors coordination when UVDE is bound to CPD or 6-4PP. Based on this finding and the outcome of the metal substitution experiments in Chapter 5 we

propose a model where the full occupation of the UVDE metal binding sites occurs only upon DNA binding and the DNA itself has an important role in the metal ions coordination.

In the crystal structure of the *T. thermophilus* UVDE an unexplained density is observed at the tip of Lys229. In Chapter 3 we give evidence that the UVDE post-translational modification is most likely carboxylation. A role for the carboxyl group is proposed in assisting the stable metal binding by donating a negative charge to the *T. thermophilus* metal coordinator His231. Lys229 is conserved in the eukaryotic UVDEs, suggesting that they might share the same modification as the *T. thermophilus* homologue. Some of the bacterial UVDEs, however, have a leucine, isoleucine, methionine, threonine or valine at the position corresponding to Lys229, thus these homologues cannot be carboxylated.

The crystal structure of UVDE gives an idea how DNA might be bound and cut by this enzyme. Initially a computer model based on the extensive structural similarities between UVDE and Endo IV suggested that UVDE most likely kinks the DNA at the site of the lesion with its GQY-loop stabilizing the kinked conformation (Chapter 2). Furthermore, in order to position the scissile phosphodiester bond close to the deeply buried metal cofactors, UVDE needs to flip the DNA damage within its active site for extrahelical repair. The initially proposed model (Chapter 2) is further elucidated in Chapter 4, where it is revealed that the initial UVDE damage recognition likely includes probing for differences of the DNA bendability. It is shown that not only a CPD, a 6-4PP or an abasic site but also a single strand nick or a gap significantly increase the binding affinity of UVDE, although not all complexes formed are catalytically active. In that regard any alteration in the DNA that increases its flexibility is a substrate for UVDE. This finding explains for the first time the previously observed broad substrate specificity of *S. pombe* UVDE. In addition, using a 2-aminopurine fluorescent probe we reveal that UVDE causes a significant destabilization in the DNA strand opposite the damage.

In the crystal structure of *T. thermophilus* UVDE Tyr105 is seen in an unusual orientation, pointing straight into the solvent. In the UVDE structural homologue Endo IV, the side-chain residue Tyr72 shows the same orientation in the absence of a DNA substrate and in the co-crystal of Endo IV with an abasic site containing DNA it projects into the kink of the DNA duplex. Based on this similarity and the fact that the alanine substitution of Tyr105 in the UVDE from *T. thermophilus* abrogates activity, we propose (Chapter 2) that this residue might be responsible for the protein-DNA interactions. The role of the tyrosine is discussed in further details in Chapter 4, where the phenotype of the *S. pombe* Y358A mutant is studied in more details. Although inactive on the abasic site lesion this mutant is highly potent in nicking the CPD and the 6-4PP lesions. We show that in the absence of Tyr358 *S. pombe* UVDE exhibits an increased catalytic activity on UV-induced lesions, but only at a lower pH of 6.5. At physiological conditions (pH 7.5) the mutant protein completely loses its catalytic activity, although it can still bind to the DNA. Based on these findings we propose that in addition to stabilizing the bend in the DNA the hydrophobic side chain of Tyr358 has a role in shielding the active site from solvent exposure.

Although originally claimed that UVDE is an alternative repair pathway for UV lesions, abasic sites, mismatches and oxidative damages we clearly demonstrated that this is not true for all UVDE homologues. We show in Chapter 4 that *S. pombe* UVDE incises DNA containing a single-strand nick or gap, but that the enzymatic activity on these substrates as well as on abasic sites strongly depends on the presence of a neighbouring pyrimidine residue. This indicates that although UVDE may have been derived from an ancestral AP endonuclease its major substrates are UV lesions and not an AP site. Furthermore, the bacterial UVDE homologues do not share the broad substrate specificity of the eukaryotic *S. pombe* protein and the enzyme from *T. thermophilus* functions as a UV damage specific nuclease.

## SAMENVATTING EN ALGEMENE DISCUSSIE

UVDE is een eiwit dat bijdraagt aan het herstel van beschadigd DNA. UVDE werd als eerste gevonden in gistsoort *Schizosaccharomyces pombe*. De eerste biochemische proeven met UVDE lieten zien dat het niet alleen actief is op door UV-licht beschadigd DNA, maar ook op DNA zonder base ('abasic sites') en DNA met 'foute' baseparen ('mismatches'). Het exacte mechanisme van UVDE voor zowel herkenning van DNA schades als voor het knippen (incisie) van beschadigd DNA was echter nog niet bekend. Van *S. pombe* UVDE was weliswaar bekend dat het  $Mn^{2+}$  en  $Mg^{2+}$  als cofactoren nodig heeft voor zijn functie, maar voor het gebruik van deze tweewaardige metaalionen bestond geen verklaring, en ook voor het aantal metaalionen dat UVDE zou kunnen gebruiken en de functie ervan bestond geen model.

Homologen van UVDE werden gevonden in vele soorten schimmels en een aantal bacteriën, zoals *Bacillus subtilis* en *Thermus thermophilus*. Hoewel de activiteit van *S. pombe* UVDE voor een deel bekend was, bestond er geen data betreffende welke DNA schades door de bacteriële UVDE-homologen herkend worden en welke metaalionen ze voor hun functie kunnen gebruiken.

De eerste weergave van UVDE, op moleculair niveau, is de kristal structuur van UVDE uit *T. thermophilus* (Hoofdstuk 2). Het UVDE enzym is een enkel TIM-barrel domein met een 29 Å brede groeve. Op de bodem van deze groeve bevindt zich een ruimte waar drie metaalionen zijn gecoördineerd. Behalve aminozuur His224, zijn alle aminozuren die betrokken zijn bij coördinatie van de metaalionen geconserveerd in alle bekende UVDE homologen. Punt mutaties in zowel *T. thermophilus* UVDE (Hoofdstuk 2) en in *S. pombe* UVDE (Hoofdstuk 5) laten zien dat coördinatie van metaal ionen van essentieel belang is voor de activiteit van UVDE, omdat de activiteit van UVDE verloren gaat na het omzetten van geconserveerde, bij metaalion coördinatie betrokken, aminozuren in alanines. De coördinatie van de drie metaalionen in UVDE is vergelijkbaar met coördinatie van de drie  $Zn^{2+}$  ionen in Endo IV, en door Endo IV als voorbeeld te nemen kan de functie van de metaalionen in UVDE, bij het knippen van DNA, begrepen worden. Hiervoor is een model opgesteld (Hoofdstukken 2 en 5), waarbij UVDE de metaalionen op positie 1 en (waarschijnlijk) positie 2 gebruikt om een nucleofiel hydroxide te maken; dit nucleofiel kan met dan de fosfodiëster banden in het DNA reageren, waarbij een instabiel intermediair gevormd wordt. Metaalion 3 zal vooral betrokken zijn bij het stabiliseren van dit instabiele intermediair.

Hoewel de fluorescentie scan geen uitsluitsel kon geven over de exacte identiteit van de gebonden metaalionen, laten biochemische experimenten zien dat het waarschijnlijk  $Mn^{2+}$  ionen zijn (Hoofdstuk 2). Metaal substitutie testen bevestigen dat inderdaad  $Mn^{2+}$  de cofactor is voor UVDE (Hoofdstuk 5). Hoewel  $Mn^{2+}$  niet het enige metaal ion is dat UVDE kan gebruiken, leidt het toevoegen van andere, tweewaardige, metaalionen tot een lagere activiteit, die ook meer afhankelijk is van het soort DNA schade. Tevens leidt het gebruik van andere metaalionen tot tegenovergestelde effecten op verschillende UVDE enzymen.  $Mg^{2+}$  heeft weliswaar een positief effect op de activiteit van *S. pombe* en *B. subtilis* UVDE, door het binden aan onbeschadigd

DNA te verlagen, maar *T. thermophilus* heeft, in aanwezig van  $Mg^{2+}$ , geen activiteit op de CPD DNA schade. Dit onverwachte effect van  $Mg^{2+}$  hangt duidelijk samen met de verminderde stabiliteit van het UVDE complex op de CPD schade. Op de 6-4PP DNA schade kan het negatieve effect van  $Mg^{2+}$  gecompenseerd worden door  $Mn^{2+}$ , omdat in aanwezigheid van beide metaalionen *T. thermophilus* UVDE efficiënt aan DNA kan binden. Deze gegevens laten zien dat, wanneer UVDE gebonden is aan CPD danwel 6-4PP, er significante verschillen bestaan in de coördinatie van metaalionen. Op basis van deze conclusie en de resultaten van metal substitutie experimenten (Hoofdstuk 5), stellen wij een model voor waarbij volledige bezetting van de metaal bindingssites in UVDE alleen voorkomt bij DNA binding en dat DNA een belangrijke rol speelt in de coördinatie van metaalionen door UVDE.

In de kristalstructuur van *T. thermophilus* UVDE werd een extra electronendichtheid gevonden aan de aminogroep van aminozuur Lys229. In Hoofdstuk 3 geven we bewijs voor een post-translationele modificatie van dit aminozuur, die zeer waarschijnlijk een carboxyl-groep is. Deze carboxyl-groep heeft een functie in metaalion coördinatie door een negatieve lading te doneren aan het, in *T. thermophilus* UVDE, metaal coördinerende aminozuur His231. Lys229 is geconserveerd in eukaryote UVDE genen, wat de suggestie geeft dat dit aminozuur in eukaryote UVDE op dezelfde manier gemodificeerd kan worden als in *T. thermophilus* UVDE. Andere, bacteriële, UVDE genen hebben een leucine, isoleucine, methionine, threonine of valine op positie van Lys229, deze aminozuur kunnen niet gemodificeerd worden met een carboxyl-groep.

De kristalstructuur van UVDE geeft een verklaring voor de manier waarop UVDE DNA schades zou kunnen binden en vervolgens knippen. Een computer gegenereerd model, gebaseerd op de overeenkomsten in de structuur van UVDE en Endo IV, liet zien dat UVDE DNA zou kunnen buigen op de plaats van een DNA schade, waarbij de GQY-loop in UVDE het gebogen DNA zou kunnen stabiliseren (Hoofdstuk 2). Om de, te knippen, fosfodiësterband van het beschadigde DNA bij de metaalionen te kunnen positioneren, moet UVDE een base flip mechanisme gebruiken om de DNA schade in de actieve site van UVDE te plaatsen. Dit oorspronkelijke model (Hoofdstuk 2) wordt verder uitgewerkt in Hoofdstuk 4, waarbij we laten zien dat de eerste stap van herkenning van DNA schades waarschijnlijk bestaat uit het zoeken naar verschillen in persistentie (de mogelijkheid tot buigen) van het DNA. Tevens laten we zien dat niet alleen CPD en 6-4PP DNA schades of DNA met een ontbrekende base, maar ook gaten in de DNA helix ('gaps') en het ontbreken van een fosfodiëster band in de DNA helix ('nicks') worden herkend door UVDE. Dit laat zien dat elke, en dus niet alleen een door UV-licht veroorzaakte, verandering in het DNA, die de lokale flexibiliteit van het DNA verhoogt door UVDE herkend en hersteld kan worden. Deze resultaten verklaren, als eerste, waarom *S. pombe* UVDE op zo veel verschillende schades actief kan zijn. Tevens, laten we, met 2-aminopurine fluorescentie bepalingen, zien dat UVDE een duidelijke destabilisatie veroorzaakt in het DNA tegenover de DNA schade.

De kristalstructuur van *T. thermophilus* UVDE laat zien dat aminozuur Tyr105 een ongebruikelijke, naar buiten gerichte, oriëntatie heeft. In Endo IV, dat wat structuur betreft een grote overeenkomst vertoont met UVDE, heeft aminozuur Tyr72 dezelfde oriëntatie. In een co-kristal van Endo IV met DNA met een abasic site, heeft aminozuur Tyr72 contact met, door Endo IV, gebogen DNA. Gebaseerd op de overeenkomsten in de structuur van Endo IV en UVDE en het feit dat het veranderen van UVDE Tyr105 in alanine UVDE inactief maakt, nemen wij aan dat aminozuur Tyr105 contact met DNA mogelijk maakt (Hoofdstuk 2). De functie van dit aminozuur wordt verder uitgewerkt in Hoofdstuk 4, waar de fenotype van de *S. pombe* Y358A in meer details is bestudeert. De substitutie van *S. pombe* Tyr358 met alanine leidt tot verlies van activiteit op abasic sites, maar niet tot verlies van activiteit op CPD en 6-4PP. Zonder Tyr358 heeft *S. pombe* UVDE een verhoogde activiteit op UV-schades (zoals CPD en (6-4)PP), maar heeft die alleen bij een (lagere) pH van 6.5. Bij een meer fysiologische pH (7.5) heeft *S. pombe* geen incisie activiteit meer zonder Tyr358, maar kan het nog wel aan DNA binden. Op basis van deze resultaten nemen we aan dat behalve het stabiliseren van gebogen DNA, Tyr358 ook een functie heeft in het afschermen van de incisie site.

Hoewel oorspronkelijk beweerd werd dat het UVDE enzym een alternatief herstelmechanisme is voor UV beschadigd DNA, abasic sites, DNA mismatches en door oxidatie reacties beschadigd DNA, laten wij zien dat dit niet het geval is voor alle UVDE homologen. In hoofdstuk 4 laten we zien dat *S. pombe* UVDE weliswaar DNA kan knippen dat een nick of gap bevat, maar dat deze activiteit op deze substraten (en ook op abasic sites) sterk afhangt van de aanwezigheid van een pyrimidine base naast de DNA schade. Dit geeft aan dat, hoewel UVDE geëvolueerd kan zijn uit een oorspronkelijk AP (abasic site) endonuclease, het eigenlijke substraat voor UVDE UV-beschadigd DNA is en niet abasic sites. Tevens is *S. pombe* UVDE actief op substraten waarop bacteriële UVDE homologen niet actief zijn en is *T. thermophilus* UVDE zelfs alleen actief op UV beschadigd DNA.

## CURRICULUM VITAE

Keti Paspaleva was born on 29 August 1978 in Bourgas, Bulgaria. In 1997 she completed the English Language School in Bourgas, Bulgaria. Between 1997 and 2003 she studied Pharmacy at the Medical University of Sofia. After a thesis project regarding the side effects of Paracetamol, she received *cum laude* a Master Degree from the Medical University of Sofia in 2003.

In the same year (2003) she started a PhD project on the structural and biochemical characterization of UVDE at the departments of Molecular Genetics and Biophysical Structural Chemistry, University of Leiden. The results of her scientific work are summarized in this thesis.

Currently she is working as a QO project leader at Centocor BV in Leiden.

## LIST OF PUBLICATIONS

Paspaleva K, Thomassen E, Pannu NS, Iwai S, Moolenaar GF, Goosen N, Abrahams JP, Crystal structure of the DNA repair enzyme ultraviolet damage endonuclease, *Structure* **15** (2007), pp. 1316-1324

Paspaleva K, Moolenaar GF, Goosen N, Damage recognition by UV damage endonuclease from *Schizosaccharomyces pombe*. *DNA Repair* (2009), in press

Meulenbroek E, Paspaleva K, Thomassen E, Abrahams JP, Goosen N, Pannu NS, Involvement of a carboxylated lysine in UV damage endonuclease from *Thermus thermophilus*, *Protein Science* (2009), in press

Paspaleva K, Moolenaar GF, Goosen N, Active site organisation of UVDE - a Mn<sup>2+</sup> dependent nuclease, *submitted*

## ACKNOWLEDGEMENTS

I would like to end this thesis by expressing my gratitude towards all colleagues who contributed to the success of this research and supported me in the last four years.

I am deeply thankful to my supervisor and co-promotor Dr. Nora Goosen for the coaching, support and encouragement. She taught me the “magic” of protein mutagenesis and helped me look at the DNA repair in a more dynamic way. She gave me the opportunity to learn, explore and apply different techniques from protein crystallization to incision assays and filter bindings.

I am very much indebted to Geri Moolenaar for her advice and incredible patience when teaching me how to work with radioactively labelled DNA. Geri’s help in solving scientific and technical problems was indispensable. I am thankful for her guidance and the fruitful discussions.

I want to thank my colleagues Erik Malta and Koen Wagner for their support, friendship and all the nice moments we had together.

I would like to thank Marian for all the practical help throughout my PhD.

I would like to express my gratitude to my promotor Prof. Jan Pieter Abrahams for the great opportunity to work in two departments: Molecular Genetics and Biophysical Structural Chemistry.

I am particularly thankful to Dr. Ellen Thomassen for the tremendous help in solving the UVDE structure. I really enjoyed working with her.

I would like to thank Dr. Harry Vrieling for all the suggestions and for reading my manuscript.

Special thanks to my friends Patrick van Eijk and Ivo Handriks, who really cheered me up in the last year of my PhD and were there every time I needed a shoulder to lean on. I would like to thank my Bulgarian friends: Stela, Pesho, Ivo Stoykov, Mitko, Diliانا and Lidiya who kept believing in me. I want to thank also Corné, Rogier and Constantijn for solving all my computer problems and being great friends.

And finally I want to thank my family who always believed in me and loved me unconditionally.

To those whose names I forgot to mention I apologize.

Keti Paspaleva

Leiden, 2009