Article No. jmbi.1999.3369 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2000) 295, 479–488

JMB

Active-site Determination of a Pyrimidine Dimer Glycosylase

John F. Garvish[1,2] and R. Stephen Lloyd[2]*

[1]*Department of Microbiology and Immunology and*

[2]*Center for Molecular Science University of Texas Medical Branch, Galveston, TX 77555-1071, USA*

**Corresponding author*

One mechanism for the repair of UV-induced DNA damage is the base excision repair pathway. The initial step in this pathway and the specificity for the type of damage that is to be repaired reside in DNA glycosylase/abasic (AP) lyases. Cleavage of the glycosyl bond of the 5′ pyrimidine of a cyclobutane pyrimidine dimer is hypothesized to occur through the destabilization of the glycosyl bond by protonation of the base or sugar with a concomitant nucleophilic attack on C1′ of the deoxyribose moiety. Based on mechanistic biochemical information from several glycosylase/AP lyases and the structural information on the bacteriophage T4 pyrimidine dimer glycosylase (T4-pdg), the catalytic mechanism has been investigated for the Chlorella virus pyrimidine dimer glycosylase (cv-pdg). As predicted from modeling studies and reaction mechanisms, the primary amine that initiates the nucleophilic displacement reaction could be trapped as a covalent imine intermediate and its identity determined by sequential Edman degradation. The primary amine was identified as the α-amino group on the N-terminal Thr2. Site-directed mutagenesis was subsequently used to confirm the conclusions that the α-amino group of cv-pdg is the active-site nucleophile.

© 2000 Academic Press

Keywords: glycosylase; DNA repair; cyclobutane pyrimidine dimer; Chlorella virus pyrimidine dimer glycosylase (cv-pdg); borohydride trapping

Introduction

The genetic integrity of cells is vulnerable to attack from a multitude of sources. One such damaging agent is UV light, which can cause two primary types of nucleotide damage, 6-4 photoproducts and cyclobutane pyrimidine dimers (Wang, 1976). In order to combat the potentially deleterious effects of this type of DNA damage, cells have evolved multiple repair mechanisms, one of which is the base excision repair (BER) pathway. This repair pathway is initiated by the class of enzymes known as glycosylase/abasic (AP) lyases (McCullough *et al.*, 1999). The most extensively characterized UV-damage specific glycosylase is the bacteriophage T4 pyrimidine dimer glycosylase (T4-pdg), previously known as endonuclease V (reviewed by Lloyd, 1998a,b). Since the structures of both T4-pdg and T4-pdg:DNA have been determined (Morikawa *et al.*, 1992, 1995) and a reaction pathway proposed (Dodson *et al.*, 1993; Fuxreiter *et al.*, 1999), this enzyme serves as a good template for structure function analyses of homologous enzymes. An enzyme available for such study is the Chlorella virus pyrimidine dimer glycosylase (cv-pdg), which functions to incise DNA at the site of cyclobutane pyrimidine dimers in double-stranded DNA (Garvish & Lloyd, 1999; McCullough *et al.*, 1998).

The design of this study was to examine the mechanism by which cv-pdg initiates the removal of DNA damage. Toward this end, and based on a 41 % amino acid identity with T4-pdg (Furuta *et al.*, 1997), molecular modeling studies have been conducted between cv-pdg and T4-pdg in order to develop a model of the structure of cv-pdg bound to its substrate DNA. This modeled structure served as a template on which to evaluate the broad base of information involving the general mechanism by which glycosylase/AP lyases cata-

Abbreviations used: AP, abasic; T4-pdg, T4 bacteriophage pyrimidine dimer glycosylase; CV-pdg, chlorella virus pyrimidine dimer glycosylase; amp, ampicillin; CS, *cis-syn*; TSII, *trans-syn* II; S, substrate; Sn, nicked substrate; BER, base excision repair.

E-mail address of the corresponding author: rslloyd@utmb.edu

 © 2000 Academic Press

lyze glycosyl and phosphodiester bond scission (reviewed by McCullough *et al*., 1999; Mol *et al*., 1999). These data suggest that there are two residues that coordinate the reactions of the glycosylase and lyase reactions. First, an acidic amino acid residue is involved in the protonation of either the endocyclic oxygen atom on the deoxyribose sugar of the damaged base or on the nitrogenous base itself; this functions to destabilize the N-C1′ glycosidic bond connecting the damaged base to the deoxyribose sugar. Alternatively, molecular dynamic analyses may suggest that actual proton transfer may be unnecessary for sufficient destabilization of the glycosyl bond (Fuxreiter *et al*., 1999). The destabilization of the glycosidic bond functions to prime C1′ of the deoxyribose sugar for nucleophilic attack. For glycosylase/AP lyases, the identity of this residue is either the α-amino group of the N terminus of the enzyme or an α-amino group of a lysine residue. Based on this information about the general reaction mechanism of glycosylase/AP lyases, as well as the structural information based on the model generated from the homology with T4-pdg, we investigated the identity of the active-site residues that are necessary for cv-pdg to catalyze glycosyl and phosphodiester bond scission on cyclobutane pyrimidine dimer-containing DNA.

Results

Construction of the modeled cv-pdg structure

The amino acid sequence of T4-pdg was aligned with the predicted amino acid sequence of the cv-pdg protein using the Look computer modeling package SegMod (Molecular Applications Group, Palo Alto, CA). The SegMod algorithm can be briefly described as follows. The cv-pdg sequence was divided into short segments, and these segments were matched to fragments in the Brookhaven Protein Databank structural database. The fragments were combined into a framework based on a template structure, in this case, the known structure of T4-pdg (Morikawa *et al*., 1995). Multiple structures of the model were built and averaged into a final structure, that was then stereochemically refined using 500 rounds of energy minimization. The RMS deviation of backbone atoms relative to the T4-pdg structure was 1.6 Å. The modeled structure of cv-pdg bound to cyclobutane pyrimidine dimer-containing DNA is shown in Figure 1(a) and an enlarged view of the predicted active-site is shown in Figure 1(b).

Covalent trapping of cv-pdg on *cis-syn* cyclobutane pyrimidine dimer and AP-site containing DNA and determination of the active-site nucleophile

Previously, it has been determined that reactions between glycosylase/AP lyases and substrate-containing DNAs that are performed in the presence of a strong reducing agent (i.e. $NaBH_4$), resulted in covalent protein-DNA complexes, suggesting a reaction mechanism (Figure 1(c)) that proceeds through the formation of a Schiff base and a β-elimination reaction (Dodson *et al*., 1993; Schrock & Lloyd, 1991). As shown in Figure 2, when 175 nM cv-pdg and 100 mM $NaBH_4$ were added to reactions containing either 0.5 nM cyclobutane pyrimidine dimer-containing DNA (left panel, lane 3) or 0.5 nM AP site-containing DNA (right panel, lane 3), cv-pdg was trapped as a covalent, dead-end complex that is stable under denaturing conditions. This complex is evident as the significantly reduced mobility of the covalent protein DNA complex (E-S for Enzyme-Substrate complex) as compared to the DNA alone (lane 1, S). In contrast, addition of enzyme with 100 mM NaCl instead of $NaBH_4$ resulted in the cleavage of both substrates, generating a ~20-mer product band (Sn) (lane 2 in both panels).

In order to determine which amino acid was covalently trapped to radiolabeled AP-containing DNA, reactions were scaled up to allow for sequence analysis of the trapped complex. The rationale behind this experiment was that if the active-site nucleophile were any of the ε-NH_2 groups in the enzyme, then N-terminal amino acid sequencing of the enzyme within the complex should yield a sequence that is identical with the unreacted enzyme, but if the α-NH_2 was the active-site nucleophile, then the Edman degradation should show evidence of an N-terminal modification. The results of sequencing cv-pdg trapped to radiolabeled AP-DNA gave results that were consistent with an N-terminal modification in two respects (data not shown). The first characteristic that suggested an N-terminal modification was the inability to definitively determine the identity of the amino acid residue for the first cycle, while the subsequent degradation cycles yielded normal amounts and mobilities consistent with the predicted second and third residues, arginine and valine, respectively. Additionally, the majority of the radio-labeled DNA was found in the first Edman degradation cycle, strongly suggesting that the N-terminal amino acid residue was covalently attached to the DNA (data not shown). The second factor indicative of an N-terminal modification was the preview of the subsequent amino acid residue for each cycle (Jue & Hale, 1994). The observation of previews is common with N-terminal modifications and consist of a phenomenon called a small peak in absorbance at the elution time of the amino acid that will appear in the following cycle. Since Met1 is cleaved post-translationally, Arg3 previewed in cycle 1, Val4 previewed in cycle 2, and Asn5 previewed in the third cycle (data not shown). Collectively, these data are strongly indicative that the N-terminal Thr2 residue is the active-site nucleophile for cv-pdg.

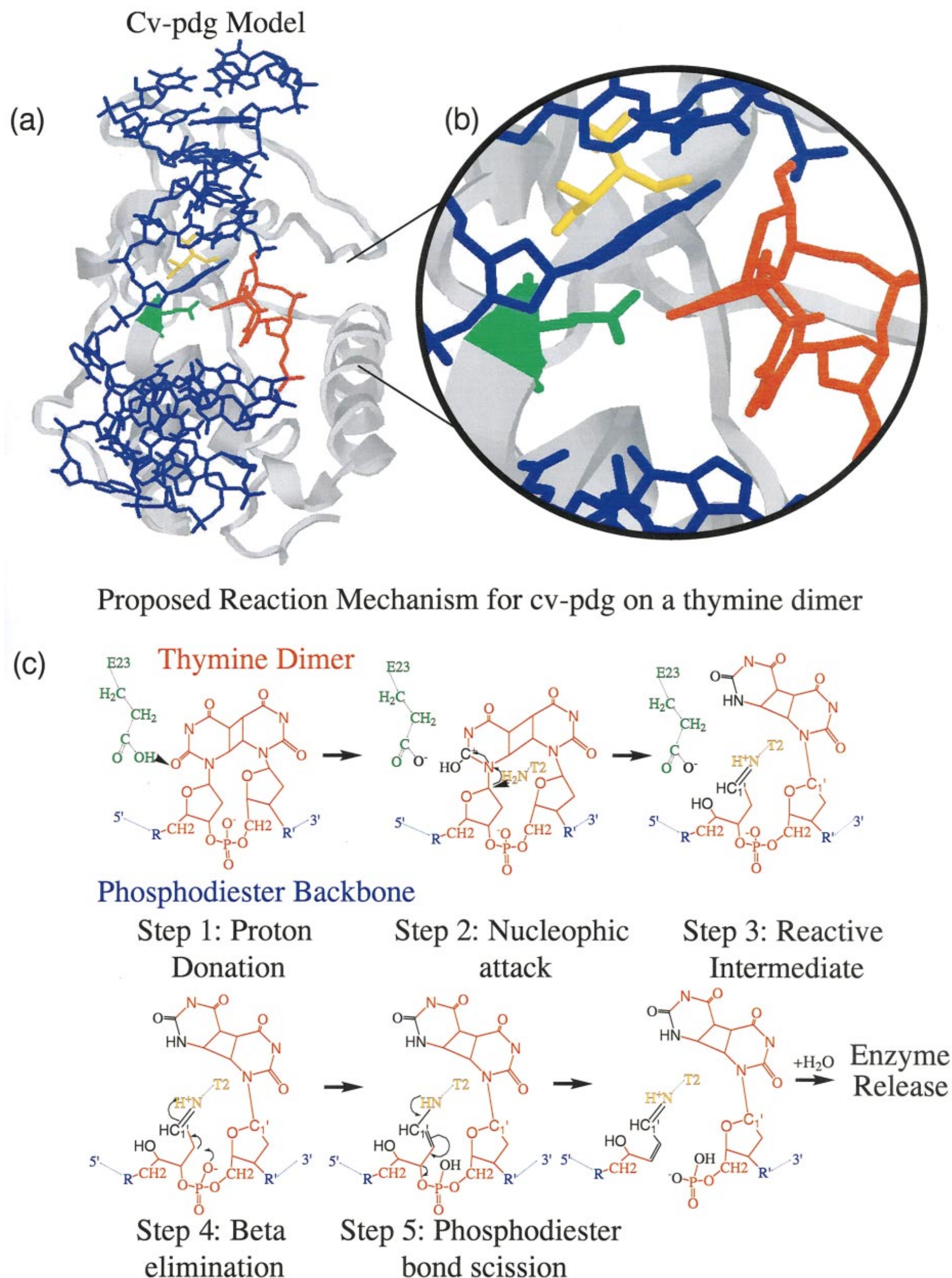

Figure 1. Chlorella virus pyrimidine dimer glycosylase modeled structure with a proposed active-site and catalytic mechanism. The modeled structure of cv-pdg is shown in (a) with the DNA represented in blue, the enzyme in gray, the *cis-syn* cyclobutane pyrimidine dimer in red, and the two proposed catalytic residues of cv-pdg Glu23 and Thr2 in green and yellow, respectively. (b) An enlarged view of the active-site of cv-pdg with the same color scheme. (c) The proposed reaction mechanism where the Glu23 (green) protonates either the thymine dimer or the endocyclic oxygen atom of the deoxyribose sugar in order to destabilize the glycosidic bond. This interaction primes C1′ of the deoxyribose sugar for a nucleophilic attack by the N-terminal amino group associated with Thr2 (yellow). The nucleophilic displacement reaction that ensues causes the formation of a transient imino intermediate whose resolution leads to the phosphodiester bond scission. This step initiates the repair by the other enzymes involved in the base excision repair pathway.

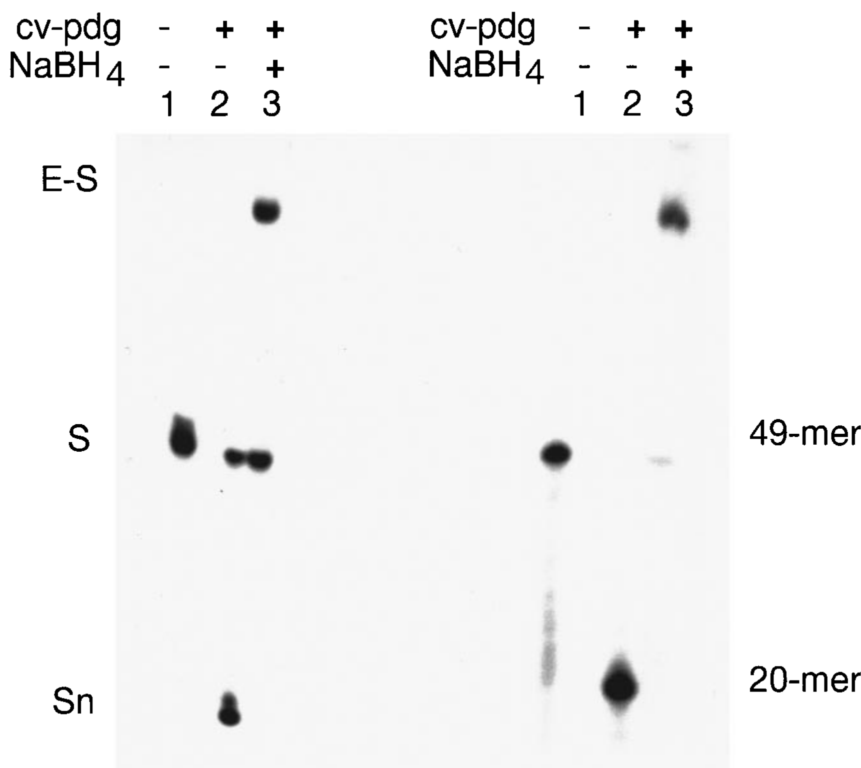

Figure 2. Formation of cv-pdg-DNA trapped complexes on cyclobutane pyrimidine dimer and AP-containing DNAs DNAs containing either a centrally located *cis-syn* thymine dimer (0.5 nM) or an abasic site (0.5 nM) were incubated with 115 nM cv-pdg. The enzyme was added to the DNA either alone (lane 2, left and right, respectively) or simultaneously with 100 mM $NaBH_4$ (lane 3, left and right, respectively). Reactions were incubated for 30 minutes at 37 °C or 25 °C for dimer or AP-containing DNA, respectively. S represents the substrate DNA band, Sn represents the nicked substrate (product) band, and the enzyme trapped to the substrate is represented by E-S.

Generation and purification of T2S, T2S/E23D, T2P, and T2P/E23D mutants of cv-pdg

In order to further investigate the role of the α-NH_2 group of Thr2 in cv-pdg, site-directed mutagenesis was carried out to introduce a codon change at Thr2 to serine (T2S) or to proline (T2P) in the pTYB2 vector containing the wild-type cv-pdg. A previously generated E23D mutation in cv-pdg was also used to generate double mutants of both the T2S/E23D and the T2P/E23D (Garvish & Lloyd, 1999). The mutant *cv-pdg* genes were sequenced to verify the single or double mutation. Following the expression of the wild-type and mutant genes in repair-deficient *Escherichia coli* (*uvrA*$^-$ *recA*$^-$), when cells were challenged with UV irradiation, only the T2S mutant displayed UV-survival enhancement levels comparable to cells expressing the wild-type enzyme; none of the other mutants could enhance survival above that observed in cells not expressing cv-pdg (data not shown). These plasmids were subsequently transformed into *E. coli* ER2566 for expression and mutant proteins expressed, and purified to homogeneity. The identity of the first 25 residues for each of these proteins was confirmed by sequential Edman degradation (data not shown).

Nicking activities on dimer and AP site-containing DNAs

In order to analyze the ability of the cv-pdg mutants to incise DNA at *cis-syn* cyclobutane pyrimidine dimers, *in vitro* incision assays were performed. Nine concentrations of cv-pdg, T2S cv-pdg, T2S/E23D cv-pdg, and T2P cv-pdg were assessed for activity on the *cis-syn* isomer of the thymine dimer, located in a 49-mer and annealed to its complementary strand (Figure 3(a)-(e)). In order to monitor every glycosylase event and not only the combined glycosylase AP lyase activities, the reactions were treated with piperidine prior to the electrophoretic separation of the substrate (S) and the nicked substrate product (Sn) bands. Figure 3(a) is a control lane of the dimer-containing DNA that was not subjected to enzyme treatment. Figure 3(b)-(e) show sequential threefold dilutions of the enzymes. These data demonstrate that the T2S cv-pdg mutant (c) had approximately fivefold less activity on *cis-syn* cyclobutane dimer-containing DNA than did the wild-type cv-pdg (b). The T2S/E23D cv-pdg double mutant (d) and the T2P cv-pdg mutant (e) both showed an approximate 100-fold reduction in activity. The T2P/E23D cv-pdg double mutant did not show any detectable activity on the *cis-syn* isomer of the thymine dimer (data not shown).

Additionally, the ability of the mutant cv-pdg to incise DNA at AP-sites was examined. In contrast to the data shown in Figure 3(b)(e), all of the enzymes appeared to have similar levels of activity on AP site-containing DNA, as can be seen from the formation of nicked substrate (Sn) in lanes 6 and 7 of (g) to (k) and the lack of product in lane 5 of each of the previously mentioned panels.

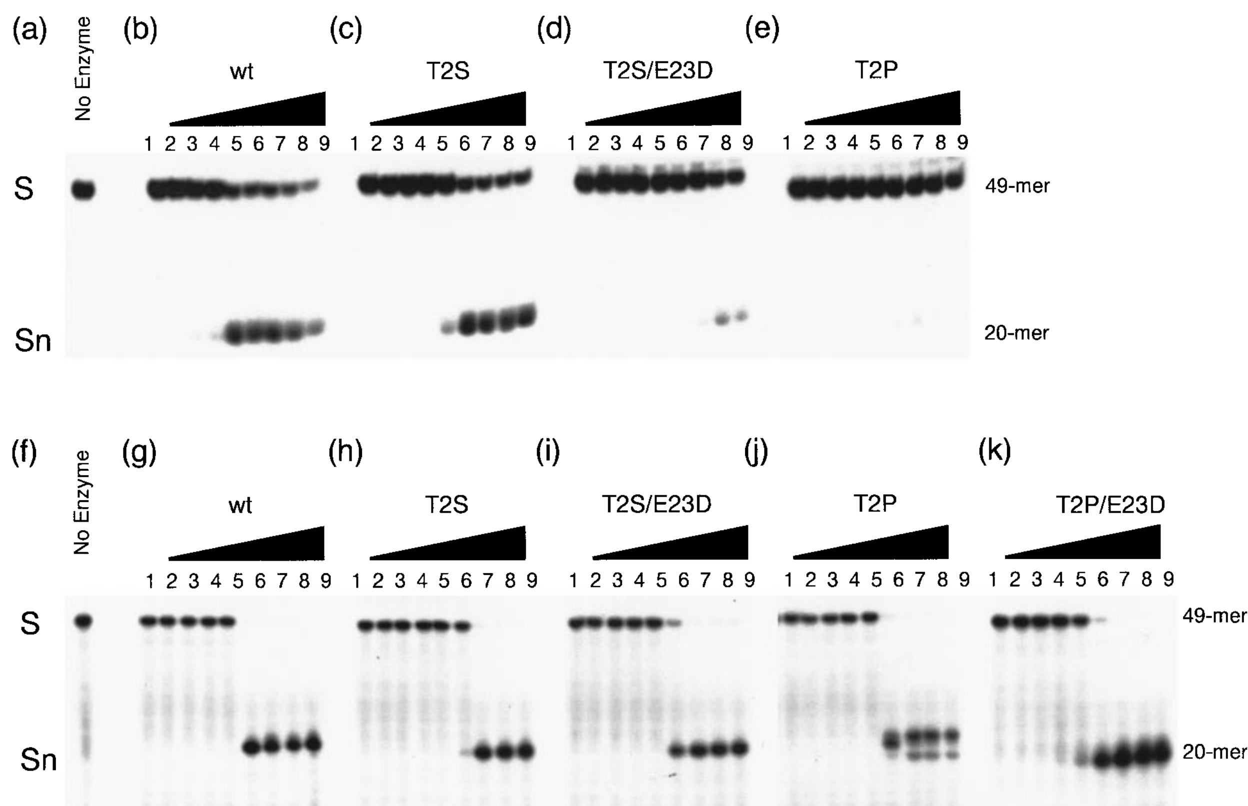

Figure 3. Incision of *cis-syn* and AP site-containing DNAs. Upper panel, an oligonucleotide containing a *cis-syn* thymine dimer (0.5 nM) was reacted with increasing concentrations of (a) no enzyme, (b) wild-type cv-pdg, (c) T2S cv-pdg, (d) T2S/E23D cv-pdg, (e) T2P cv-pdg for 30 minutes at 37 °C. S represents the DNA substrate band and Sn represents the nicked-substrate (product) band. Following the 30 minute incubation, each reaction was treated with piperidine to convert any DNAs containing AP sites to the incised product band. The reaction products were separated by electrophoresis on a 15 % polyacrylamide denaturing gel containing 8 M urea. Lanes 1-9 are at 0, 0.01, 0.05, 0.15, 0.47, 1.41, 4.25, 12.77, 38.3, 115 nM for each enzyme except T2P/E23D cv-pdg, which had no detectable activity (data not shown). Lower panel, an oligonucleotide containing an abasic site (0.5 nM) was reacted with increasing concentrations of (g) wild-type cv-pdg, (h) T2S cv-pdg, (i) T2S/E23D cv-pdg, (j) T2P cv-pdg, (k) T2P/E23D cv-pdg at 25 °C. (f) No enzyme. The concentrations of the enzymes for lanes 1-9 were as above.

Covalent trapping analyses on dimer and AP site-containing DNAs

Wild-type cv-pdg and its mutants were examined for their ability to be covalently trapped on thymine dimer-containing DNA (Figure 4(a)). Both the wild-type cv-pdg and the T2S mutant showed activity on the *cis-syn* thymine dimer, as is evident from the conversion of substrate (S) 49-mer to nicked substrate (Sn) product 20-mer (lanes 2 and 4). Both of these enzymes were able to be covalently trapped in the presence of $NaBH_4$, as can be seen from the formation of the enzyme substrate covalent complex (E-S) (lanes 3 and 5). The double mutants T2S/E23D and T2P/E23D cv-pdg showed no formation of nicked substrate (lanes 6 and 10) or formation of the covalent enzyme substrate complex (lanes 7 and 11). The T2P cv-pdg mutant was able to form a small amount of nicked substrate (lane 8), although the mobility of the product band appears to be a result of a majority of δ-elimination reaction rather than the combined β/δ elimination reaction as seen for the wild-type and T2S mutant (lanes 2 and 4). The absence of the ability to form the covalent complex for T2P is probably due to the low levels of activity.

In addition, the mutant cv-pdg enzymes were examined for their ability to form the covalent enzyme substrate complex (E-S) when the reaction is carried out in the presence of $NaBH_4$ (Figure 4(b)). Lane 1 is a control lane without the addition of any enzyme. Lanes 2, 4, 6, 8, and 10 show the activity that is present on AP site-containing DNA with the addition of excess pure enzyme. The double mutants T2S/E23D and T2P/E23D (lanes 7 and 11) showed a low level of formation of the enzyme substrate (E-S) trapped complex. The wild-type and T2S mutant (lanes 3 and 5) appear to trap at relative efficiencies similar to that seen for the *cis-syn* CPD; however, the trapping appears to be more efficient for both of these enzymes on the AP site-containing DNA. However, the T2P cv-pdg (lane 9) mutant appeared to be the most readily formed enzyme substrate complex on AP site-containing DNA.

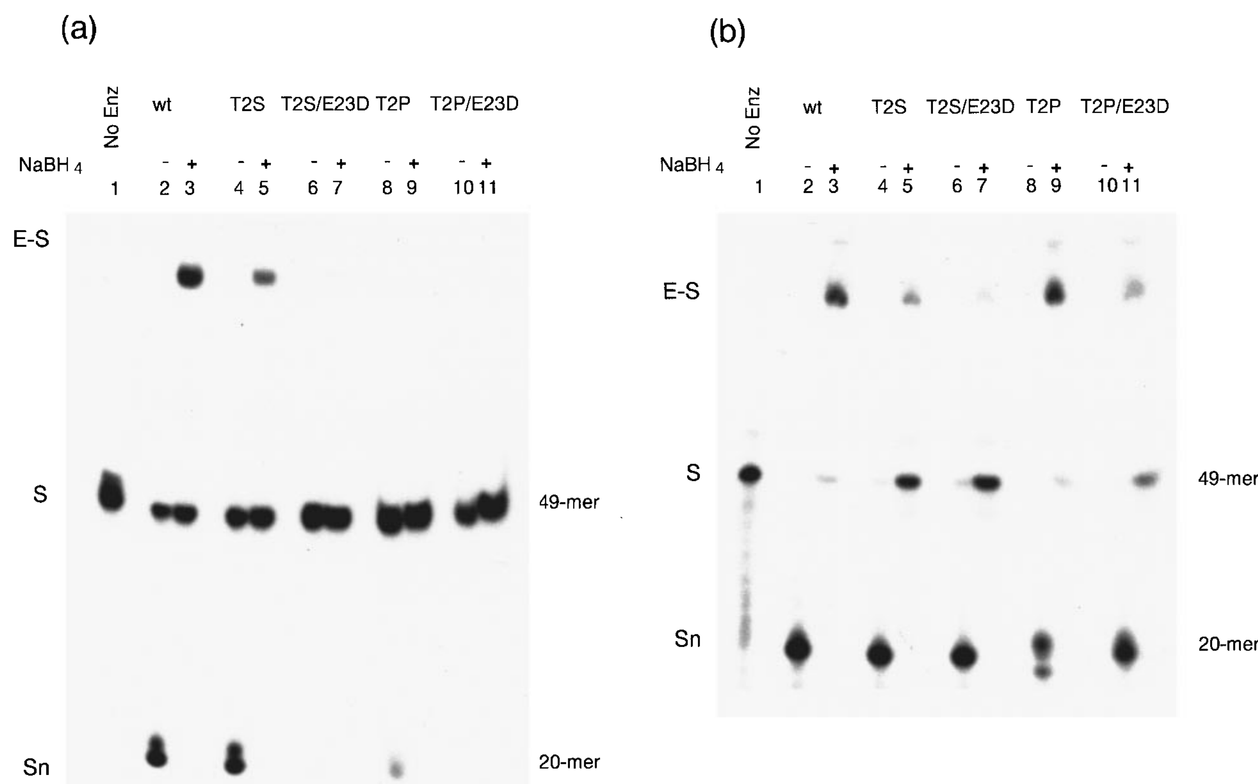

Figure 4. Covalent trapped complex formation of cv-pdg and cv-pdg mutants on *cis-syn* thymine dimer and AP site-containing DNA. DNA containing (a) a centrally located (0.5 nM) *cis-syn* thymine dimer or (b) an AP site (0.5 nM) were incubated with 115 nM cv-pdg, T2S cv-pdg, T2S/E23D cv-pdg, T2P cv-pdg, T2P/E23D cv-pdg. The enzymes were either added to the DNA alone (lanes 2, 4, 6, 8, and 10, respectively) or were added with 100 mM $NaBH_4$ simultaneous with the enzyme (lanes 3, 5, 7, 9, and 11, respectively). Reactions were incubated for 30 minutes at 25 °C. S represents the substrate DNA band, Sn represents the nicked substrate (product) band, and the enzyme trapped to the substrate is represented by E-S.

Amino acid sequencing of T2P cv-pdg trapped to radioactively labeled AP-containing DNA

Given the ability to trap the T2P cv-pdg mutant on AP-containing DNA, experiments were performed to sequence the T2P AP-DNA covalently trapped complex. These experiments are qualitatively similar to work from the Grollman laboratory involving the inability to obtain direct protein sequence data *via* Edman degradation from an N-terminal active-site proline residue of Fpg that had been covalently trapped to substrate DNA (Zharkov *et al*., 1997). The formation of the covalently trapped complex and transfer onto PVDF paper were conducted as previously described. There was no detectable Edman degradation product, suggesting one of two scenarios. The first was that there was insufficient protein for amino acid sequence determination. The second was that there was an N-terminal modification of the N-terminal proline residue. To address the first possibility, the concentration of the protein-DNA complex was determined for the samples whose sequence could not be obtained. These data revealed that there was at least 20-fold greater amount (10 pmol) of protein than necessary in the sample as compared to the minimum amount needed for sequencing. Thus, the other plausible explanation was that there was an N-terminal modification of the enzyme to the AP-DNA. To analyze this possibility, the fractions collected from the HPLC column during the T2P protein sequencing were analyzed for radioactivity and these were found to have insignificant levels. The sequence data for the wild-type cv-pdg-AP DNA complex (a primary amine at the N terminus converted to a secondary amine upon covalent trapping, and still subject to Edman degradation) and the lack of sequence data for the T2P cv-pdg-AP DNA complex (secondary amine converted to tertiary amine upon covalent trapping, no longer subject to Edman degradation), are what would be expected for an enzyme that functions using its N terminus as the active-site nucleophile. These data indicate an N-terminal active-site nucleophile for the wild-type cv-pdg in the reaction catalyzing the excision of cyclobutane pyrimidine dimers from DNA.

Discussion

With T4-pdg serving as a prototype for the analysis of active-site residues in the glycosylase/AP lyase class of enzymes, a generalized catalytic mechanism has evolved that predicts a concerted and coordinated action of two amino acid residues that appear to be essential for this first step of the base excision repair reaction (reviewed by Lloyd, 1998a,b; McCullough *et al*., 1999; Mol *et al*., 1999). Following specific damage recognition, which involves a nucleotide flipping reaction, the initial step is the destabilization of the N-C1′ glycosidic bond. This is likely to occur through the development of a positive charge on the endocyclic oxygen atom of the deoxyribose sugar without an obligatory proton transfer step (Fuxreiter *et al*., 1999). Following this glycosidic bond destabilization, C1′ of the deoxyribose sugar is primed for nucleophilic attack by either the α-NH_2 group of the N terminus or the ε-NH_2 group on lysine residues. In the glycosylase/AP lyase family, there is literature precedent for either the α-NH_2 of the N terminus or the ε-NH_2 group from a lysine residue to catalyze the nucleophilic attack on C1′: α-NH_2-T4-pdg (Schrock & Lloyd, 1991), FPG (Zharkov *et al*., 1997), ε-NH_2-*E. coli* endonuclease III (Thayer *et al*., 1995), and yeast Ogg1 (Nash *et al*., 1997). With these data serving as a foundation, we investigated the identity of the active-site nucleophile of cv-pdg. Upon reaction with DNAs containing either a cyclobutane pyrimidine dimer or an abasic site, cv-pdg could be covalently trapped on these DNAs when the reactions were carried out in the presence of the strong reducing agent, $NaBH_4$. To identify the specific amino acid that was covalently bound to these DNAs, the protein-DNA complexes were subjected to Edman degradation analyses. These data revealed that the N-terminal threonine residue migrated at an anomolous position, which coincided perfectly with the radiolabeled DNA. The conclusion from these studies was that the N-terminal threonine residue is the active-site nucleophile.

With a considerable amount of evidence implicating the N terminus of cv-pdg as the active-site nucleophile, site-directed mutations at Thr2 (T2S and T2P) were constructed. Additionally, double mutants of the N-terminal threonine and glutamic acid to aspartic acid at position 23 (T2S/E23D and T2P/E23D) were created. To obtain a biological endpoint for the effect of these mutations, the wild-type and mutant cv-pdg enzymes were examined for their ability to enhance the survival of recombinational and nucleotide excision repair-deficient *E. coli*. The wild-type cv-pdg conferred considerable survival enhancement over control cells containing vector alone and, as predicted, the T2S mutant of cv-pdg showed survival levels similar to that of the wild-type enzyme. These data are consistent with the more sensitive *in vitro* assays that showed similar activity of this mutant protein on *cis-syn* dimer and AP site-containing DNA relative to that of the wild-type enzyme. The other mutants, E23D, T2P, T2P/E23D, and T2S/E23D, showed no detectable enhancement in UV survival over cells containing the plasmid alone. Most likely, this result can be attributed to a loss in DNA glycosylase activity at *cis-syn* cyclobutane pyrimidine dimers, as evident in Figure 3.

In order to investigate the activities of these cv-pdg mutants relative to the wild-type enzyme, all enzymes were assayed using either a synthetic oligonucleotide containing a site-specific *cis-syn* thymine dimer or apurinic site. As would be predicted from the proposed reaction mechanism, the T2S cv-pdg mutant had only slightly reduced activity when compared to the wild-type enzyme, a result consistent with the proposal in which the α-NH_2 group is responsible for the activity and the exact identity of the side-chain should not be of critical importance. However, the double mutant T2S/E23D cv-pdg had an ~80-fold reduced activity on the *cis-syn* thymine dimer-containing substrate. These data are generally consistent with the additive effects of mutations at the two critical active-site residues. The T2P cv-pdg mutant had almost no glycosylase activity on *cis-syn* thymine dimers, a result suggesting that by altering the α-NH_2 group at the N terminus from a primary amine to a secondary amine, the catalytic efficiency of initiating dimer cleavage was severely compromised due to steric restrictions or alterations in the pK_a of the secondary amine. The T2P/E23D cv-pdg double mutant had no detectable activity on *cis-syn* dimer-containing DNA, a result that was not surprising, considering the results for the individual mutant proteins. However, the more flexible AP-containing DNA retained almost equivalent activity for all of the mutants. These data suggest that this or some other characteristic of the dimer-containing DNA, perhaps the bend in the DNA induced by the presence of the lesion (Yamaguchi *et al*., 1998), that is not present in the AP-containing DNA, renders it less amenable to alterations in the catalytic center of the enzyme.

A final test of the hypothesis that the N terminus of cv-pdg is the active-site nuelophile, was to covalently trap the T2P mutant on an AP site, and test whether it would no longer be susceptible to Edman degradation, since the secondary amine of proline would be converted to a tertiary amine and not be subject to this degradation chemistry. In this regard, although >10 pmol of protein-DNA complex was submitted for analysis, no detectable sequence for the T2P-DNA complex was obtained, consistent with an N-terminal blockage. Additionally, when fractions collected from the HPLC elution profile were analyzed for the presence of radioactive substrate DNA, no significant radioactivity was detected. These data provide further evidence that the DNA is trapped to the N terminus of the protein, since it cannot undergo Edman degradation, the DNA stays bound to the protein and does not elute over the HPLC column.

Collectively, these experiments establish that the α-NH_2 of cv-pdg is the active-site nucleophile responsible for the initiation of base excision reaction involved in thymine dimer repair. Taken together with the previous work on the role of Glu23 in the destabilization of the glycosidic bond (Garvish & Lloyd, 1999), this is further validation of the reaction mechanism shown in Figure 1(c).

Materials and Methods

Cloning of Chlorella virus pyrimidine dimer glycosylase

The cv-*pdg* gene was PCR amplified from the pET11a vector to generate a fragment with an *Xma*I site on the 3′ end. The PCR conditions were 95 °C for ten minutes, 25 cycles of 95 °C 30 seconds, 52 °C for 60 seconds, 68 °C for 90 seconds, followed by 68 °C for ten minutes. The primers for the PCR reaction had the following sequence: 5′ CCCGGGATTATTGCTGGTTTTAGCTTTCGTGAAC 3′, the 3′ chitin primer; and 5′ TTAACTTTAAGAAGGAGATATACATATG 3′, the 5′ chitin primer. The PCR amplified fragment was then cut with *Nde*I and *Xma*I to obtain a fragment, that was subsequently subcloned into the pTYB2 vector, a chitin-binding domain protein fusion expression vector (New England Biolabs). A ligation reaction with the cut pTYB2 vector and the cut PCR fragments was conducted for 12 hours at 16 °C. The ligation reaction (4 μl) was transformed into DH5α *E. coli* cells. The colonies were screened for pTYB2 vector with a *cv-pdg* gene insert using an oligonucleotide complementary to a sequence present on the *cv-pdg* gene; 5′ TTTTGATGGGATCTTCTTCAA 3′.

The clones that appeared positive through differential hybridization were then confirmed to be the appropriate sequence *via* automated DNA sequencing (performed by NIEHS Molecular Biology Core at the University of Texas Medical Branch under the supervision of Thomas G. Wood).

Amino acid sequencing of cv-pdg and cv-pdg-DNA trapped complex

A total of 10 μg of cv-pdg or T2P cv-pdg was added simultaneously with $NaBH_4$ to 10 μg of AP site-containing, ^{32}P-labeled or non-labeled 25-mer DNA. This reaction was incubated at 25 °C for 30 minutes and then lyophilized. The E-S trapped complex was separated from the uncomplexed enzyme and substrate by electrophoresis on an SDS/15 % polyacrylamide gel. The gel was then electroblotted onto PVDF paper (Bio-Rad) for one hour at 50 V in a buffer containing Caps, 10 % (v/v) methanol, and 0.05 % (w/v) SDS. Following the transfer, the PVDF paper was stained with Coomassie brilliant blue R-250 in 50 % methanol, 10 % acetic acid and the bands of the enzyme alone and the E-S trapped complex were excised from the paper and submitted for automated Edman degradation amino acid sequencing on a Perkin Elmer494/HT Protein Sequencer (Foster City, CA) by the NIEHS Center Protein Chemistry Core (UTMB).

Oligonucleotide site-directed mutagenesis of cv-pdg

Site-directed mutagenesis was carried out on the pTYB2 vector containing the cv-*pdg* gene using the Quikchange Mutagenesis Kit (Stratagene). The sequences of the primers used to carry out the mutagenesis of the cv-*pdg* gene were as follows: 5′ CTTTAAGAAGGAGATATACATATG**T**CACGTGTGAATCTCGTACCGGTTCAG 3′ (T2S +), 5′ CTGAACCGGTACGAGATTCACACGTG**A**CATATGTATATCTCCTTCTTAAAG 3′ (T2S −), 5′ CTTTAAGAAGGAGATATACATATG**C**CACGTGTGAATCTCGTACCGGTTCAA 3′ (T2P +), and 5′ TTGAACCGGTACGAGATTCACACGTG**G**CATATGTATATCTCCTTCTTAAAG 3′ (T2P −), where the mismatched nucleotides are in bold face. The PCR conditions were 95 °C for seven minutes, 25 cycles of 95 °C for 60 seconds, 55 °C for 60 seconds, 63 °C for 18 minutes, followed by 72 °C for seven minutes. Following the PCR reactions, the templates were digested with *Dpn*I and the PCR product visualized following separation on a 0.8 % agarose gel and staining with 1 μg/ml ethidium bromide in TAE (0.04 M Tris-acetate, 1 mM EDTA). An aliquot of the PCR reaction (4 μl) was used for transformation, following standard protocols, into Supercompetent XL1 Blue bacterial cells (Stratagene). After one hour at 0 °C, the transformation reaction was heat-shocked for one minute at 42 °C. The cells were allowed to recover for one hour at 37 °C in 2 × YT medium (pH 7.0) (16 g of Bacto-tryptone, 10 g of Bacto-yeast extract and 5 g of NaCl per liter of deionized water). Cells were then plated out on LB plates that contained 100 μg/ml ampicillin (amp). Plasmid preparations were prepared from amp-resistant colonies and the mutation was confirmed by automated sequencing (NIEHS Molecular Biology Core Facility under the direction of Dr T. G. Wood).

Purification of cv-pdg mutants T2S, T2S/E23D, T2P, and T2P/E23D

The purification scheme for the mutant enzymes was essentially as described (Furuta *et al*., 1997; Garvish & Lloyd, 1999).

Covalent trapping of imino-intermediate using sodium borohydride on dimer and AP site-containing DNA

Oligonucleotides containing either a *cis-syn* thymine dimer or an abasic site were diluted to obtain a final substrate concentration of 0.5 nM in the reaction buffer. These 49-mers containing a site specific *cis-syn* thymine dimer were provided generously by Colin Smith and John-Stephen Taylor (Washington University, St. Louis, MO) with the sequence 5′ AGCTACCATGCCTGCACGAATTAAGCAATTCGTAATCATGGTCATAGCT 3′ (McCullough *et al*., 1998; Smith & Taylor, 1993). The TT shows the position of the dimer. A 49-mer oligonucleotide containing a site-specific uracil was synthesized (Midland Research) with the sequence 5′ AGCTACCATGCCTGCACGAAUTAAGCAATTCGTAATCATGGTCATAGCT 3′. The U shows the position of the uracil base. This uracil-containing 49-mer was ^{32}P-labeled on the 5′ end with phage T4 polynucleotide kinase and annealed to its complementary oligonucleotide. The double-stranded uracil-containing 49-mer was incubated with uracil DNA glycosylase (Epicentre Technologies) for ten minutes at 37 °C, generating an AP site. These oligonucleotides were then diluted with reaction buffer (25 mM NaH_2PO_4 (pH 6.8), 1 mM EDTA, 100 mM NaCl, 100 μg/ml bovine serum albumin). Reactions were initiated by the simultaneous addition of

100 mM $NaBH_4$ or 100 mM NaCl with 115 nM cv-pdg or mutant cv-pdg. All reactions were incubated for 30 minutes at 37 °C (dimer-containing DNA) or 25 °C (for AP-containing DNA). These reactions were terminated by the addition of the loading buffer (95 % (v/v) formamide, 20 mM EDTA, 0.02 % (w/v) bromophenol blue, 0.02 % (w/v) xylene cyanol) and the trapped complexes were separated from free DNAs by electrophoresis through an 8 M urea, 15 % polyacrylamide denaturing gel. The DNA bands were visualized by autoradiography of the wet gels using Hyperfilm-MP X-ray film (Amersham).

Nicking activity of cv-pdg, and cv-pdg mutants

The same DNAs described in the trapping section were used for the activity assays. These DNAs were labeled on the 5′ end with [γ-^{32}P]ATP by T4 polynucleotide kinase and annealed to an unlabeled complementary oligonucleotide. The double-stranded thymine dimer containing 49-mer was diluted with reaction buffer (25 mM NaH_2PO_4 (pH 6.8), 1 mM EDTA, 100 mM NaCl, 100 μg/ml bovine serum albumin) and the appropriate concentrations of the cv-pdg, T2S cv-pdg, T2S/E23D cv-pdg, T2P cv-pdg, or T2P/E23D cv-pdg were added to the reactions for 30 minutes at 37 °C. Incision reactions were terminated by the addition of loading buffer. Dimer-containing DNAs were subsequently treated with 1 M piperidine and heated to 100 °C for ten minutes. The purpose of the piperidine treatment was to convert any AP sites into single-strand breaks. The substrate DNAs were separated from the incision product DNAs by electrophoresis through a denaturing 15 % polyacrylamide gel containing 8 M urea. The DNA bands were visualized by autoradiography of the wet gels using Hyperfilm-MP X-ray film (Amersham).

Acknowledgments

We thank J.S. Taylor and C. Smith (Washington University, St. Louis, MO) for their generous gift of DNAs containing site-specific *cis-syn* thymine dimers. The molecular modeling was performed in collaboration with M. L. Dodson (Sealy Center for Molecular Science), W. Braun, C. Schein, and Hongyao Zhu (Sealy Center for Structural Biology). The DNA sequencing, oligonucleotide synthesis were performed by the NIEHS Center Molecular Biology Core under the direction of Dr T. G. Wood, and the amino acid sequencing was performed in the NIEHS Center Protein Chemistry and Purification Core facility by Dr Alex Kurosky and John S. Smith. We thank Drs Amanda McCullough and Paul House for critical reading of this manuscript. This work was supported by NIEHS T327254, ES04091, and ES06676. R.S.L. holds the Mary Gibbs Jones Distinguished Chair in Environmental Toxicology from the Houston Endowment.

References

Dodson, M. L., Schrock, R. D. D. & Lloyd, R. S. (1993). Evidence for an imino intermediate in the T4 endonuclease V reaction. *Biochemistry*, **32**, 8284-8290.

Furuta, M., Schrader, J. O., Schrader, H. S., Kokjohn, T. A., Nyaga, S., McCullough, A. K., Lloyd, R. S., Burbank, D. E., Landstein, D., Lane, L. & Van Etten, J. L. (1997). Chlorella virus PBCV-1 encodes a homolog of the bacteriophage T4 UV damage repair gene denV. *Appl. Environ. Microbiol.* **63**, 1551-1556.

Fuxreiter, M., Warshel, A. & Osman, R. (1999). Role of active-site residues in the glycosylase step of T4 endonuclease V. Computer simulation studies on ionization states. *Biochemistry*, **38**, 9577-9589.

Garvish, J. F. & Lloyd, R. S. (1999). The catalytic mechanism of a pyrimidine dimer-specific glycosylase (pdg)/abasic lyase, Chlorella virus-pdg. *J. Biol. Chem.* **274**, 9786-9794.

Jue, R. A. & Hale, J. E. (1994). On-line procedures for alkylation of cysteine residues with 3-bromopropylamine prior to protein sequence analysis. *Anal. Biochem.* **221**, 374-378.

Lloyd, R. S. (1998a). Base excision repair of cyclobutane pyrimidine dimers. *Mutat. Res. DNA Repair*, **408**, 159-170.

Lloyd, R. S. (1998b). The initiation of DNA base excision repair of dipyrimidine photoproducts. *Progr. Nucl. Acid Res.* **62**, 155-175.

McCullough, A. K., Romberg, M. T., Nyaga, S., Wei, Y., Wood, T. G., Taylor, J. S., Van Etten, J. L., Dodson, M. L. & Lloyd, R. S. (1998). Characterization of a novel *cis-syn* and *trans-syn*-II pyrimidine dimer glycosylase/AP lyase from a eukaryotic algal virus, *Paramecium bursaria* chlorella virus-1. *J. Biol. Chem.* **273**, 13136-13142.

McCullough, A. K., Dodson, M. L. & Lloyd, R. S. (1999). Initiation of base excision repair: glycosylase mechanisms and structures. *Annu. Rev. Biochemistry*, **68**, 255-285.

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

Morikawa, K., Matsumoto, O., Tsujimoto, M., Katayanagi, K., Ariyoshi, M., Doi, T., Ikehara, M., Inaoka, T. & Ohtsuka, E. (1992). X-ray structure of T4 endonuclease V: an excision repair enzyme specific for a pyrimidine dimer. *Science*, **256**, 523-526.

Morikawa, K., Ariyoshi, M., Vassylyev, D. G., Matsumoto, O., Katayanagi, K. & Ohtsuka, E. (1995). Crystal structure of a pyrimidine dimer-specific excision repair enzyme from bacteriophage T4: refinement at 1.45 Å and X-ray analysis of the three active-site mutants. *J. Mol. Biol.* **249**, 360-375.

Nash, H. M., Lu, R., Lane, W. S. & Verdine, G. L. (1997). The critical active-site amine of the human 8-oxoguanine DNA glycosylase, hOgg1: direct identification, ablation and chemical reconstitution. *Chem. Biol.* **4**, 693-702.

Schrock, R. D. III & Lloyd, R. S. (1991). Reductive methylation of the amino terminus of endonuclease V eradicates catalytic activities. Evidence for an essential role of the amino terminus in the chemical mechanisms of catalysis. *J. Biol. Chem.* **266**, 17631-17639.

Smith, C. A. & Taylor, J. S. (1993). 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. *J. Biol. Chem.* **268**, 11143-11151.

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

Wang, S. Y. (1976). *Photochemistry and Photobiology of Nucleic Acids*, vol. 1, Academic Press, New York.
Yamaguchi, H., van Aalten, D. M., Pinak, M., Furukawa, A. & Osman, R. (1998). Essential dynamics of DNA containing a *cis.syn* cyclobutane thymine dimer lesion. *Nucl. Acids Res.* **26**, 1939-1946.
Zharkov, D. O., Rieger, R. A., Iden, C. R. & Grollman, A. P. (1997). NH_2-terminal proline acts as a nucleophile in the glycosylase/AP-lyase reaction catalyzed by *Escherichia coli* formamidopyrimidine-DNA glycosylase (Fpg) protein. *J. Biol. Chem.* **272**, 5335-5341.

Edited by I. Tinoco

(Received 15 July 1999; received in revised form 4 November 1999; accepted 7 November 1999)