Cell, Vol. 83, 773–782, December 1, 1995, Copyright © 1995 by Cell Press

Atomic Model of a Pyrimidine Dimer Excision Repair Enzyme Complexed with a DNA Substrate: Structural Basis for Damaged DNA Recognition

Dmitry G. Vassylyev,* Tatsuki Kashiwagi,* Yuriko Mikami,* Mariko Ariyoshi,* Shigenori Iwai,† Eiko Ohtsuka,† and Kosuke Morikawa*
*Protein Engineering Research Institute
6-2-3 Furuedai
Suita
Osaka 565
Japan
†Faculty of Pharmaceutical Sciences
Hokkaido University
Sapporo
Hokkaido 060
Japan

Summary

T4 endonuclease V is a DNA repair enzyme from bacteriophage T4 that catalyzes the first reaction step of the pyrimidine dimer-specific base excision repair pathway. The crystal structure of this enzyme complexed with a duplex DNA substrate, containing a thymine dimer, has been determined at 2.75 Å resolution. The atomic structure of the complex reveals the unique conformation of the DNA duplex, which exhibits a sharp kink with a 60° inclination at the central thymine dimer. The adenine base complementary to the 5′ side of the thymine dimer is completely flipped out of the DNA duplex and trapped in a cavity on the protein surface. These structural features allow an understanding of the catalytic mechanism and implicate a general mechanism of how other repair enzymes recognize damaged DNA duplexes.

Introduction

Radiation and some chemical agents frequently produce modified bases within DNA that prevent accurate replication or transcription. These modified bases can potentially be mutagenic or lethal for the cell, and hence organisms usually have systems that can excise modified bases from the genome and replace them with normal nucleotides. This process, termed base excision repair, can proceed through various enzymic reactions. In some cases, damaged bases are cleaved by the action of DNA glycosylases, which catalyze the cleavage of the N-glycosylic bonds between the modified bases and their sugar moieties. This reaction leaves apurinic/apyrimidinic (AP) sites within the DNA, which are then removed by AP endonucleases. These endonucleases specifically catalyze the hydrolysis of phosphodiester bonds at abasic sites in the DNA. The precise mechanism of the DNA excision repair, which is prevalent from prokaryotes to eukaryotes, remains unknown, although a full understanding of its molecular basis is crucial for cell biology and the medical sciences.

Ultraviolet irradiation causes the formation of cyclobutane-type pyrimidine dimers (PDs) within DNA. Endonuclease V (endo V) (endodeoxyribonuclease [PD], EC3.1.25.1) from bacteriophage T4 is an enzyme that catalyzes the first step of the DNA excision repair pathway. The enzyme, which consists of 137 amino acids, possesses two distinct catalytic activities: a PD-specific glycosylase and an AP lyase (Macmillan et al., 1981; Nakabeppu and Sekiguchi, 1981; Dodson et al., 1994; Latham and Lloyd, 1994). These reactions proceed through two steps: scission of the glycosyl bond at the 5′ side of the PD (Dodson et al., 1993), followed by incision of the phosphodiester bond at the 3′ position of the abasic site through β elimination to produce an α,β-unsaturated aldehyde and a 5′-terminal monophosphate as final products (Kim and Linn, 1988; Manoharan et al., 1988). This enzymatic activity is found in certain microorganisms, such as Micrococcus luteus (Sancer and Sancer, 1988) and Saccharomyces cerevisiae (Hamilton et al., 1992), and differs from the ultraviolet excision repair system consisting of the *uvr* gene products in Escherichia coli (Sancer and Sancer, 1988). The enzyme binds only to double-stranded DNA (Inaoka et al., 1989) in two different manners: the enzyme scans nontarget DNA sequences by electrostatic interactions to search for damaged sites and, subsequently, specifically recognizes the PD site (Dodson et al., 1994).

In recent years, T4 endo V was extensively studied by both biochemical methods and X-ray crystallography. The crystal structure of the enzyme was determined at 1.6 Å resolution (Morikawa et al., 1992). In combination with site-directed mutagenesis, the structural features of the enzyme allowed the identification of the catalytic center of the glycosylase (Doi et al., 1992; Morikawa et al., 1992; Morikawa, 1993). More recently, crystallographic studies of three active site mutants have provided more defined insights into the roles of two key residues (Glu-23 and Arg-3) in the glycosylase catalytic center, while the crystal structure refined at 1.45 Å resolution revealed the detailed architecture at the atomic level (Morikawa et al., 1994, 1995). On the other hand, based on the biochemical and mutational analyses combined with nuclear magnetic resonance (NMR), Lloyd and associates proposed a reaction scheme for T4 endo V in which an imino–covalent enzyme–DNA substrate intermediate is formed (Schrock and Lloyd, 1991, 1993). The structural features of the enzyme in the absence of the DNA substrate are found to be generally consistent with this proposed reaction scheme. However, structural images of the enzyme–DNA complex, such as the conformation of the bound DNA, the precise recognition mechanism of the damaged DNA, and the catalytic mechanism, have remained obscure.

We report here the cocrystal structure of T4 endo V (E23Q mutant) in complex with a DNA duplex containing a thymine dimer. The three-dimensional structure of the complex refined at 2.75 Å resolution reveals the unique structure of the bound DNA duplex, which provides detailed insights into the catalytic mechanism and implicates a general recognition mechanism by which proteins can discriminate a damaged DNA duplex from a normal one.

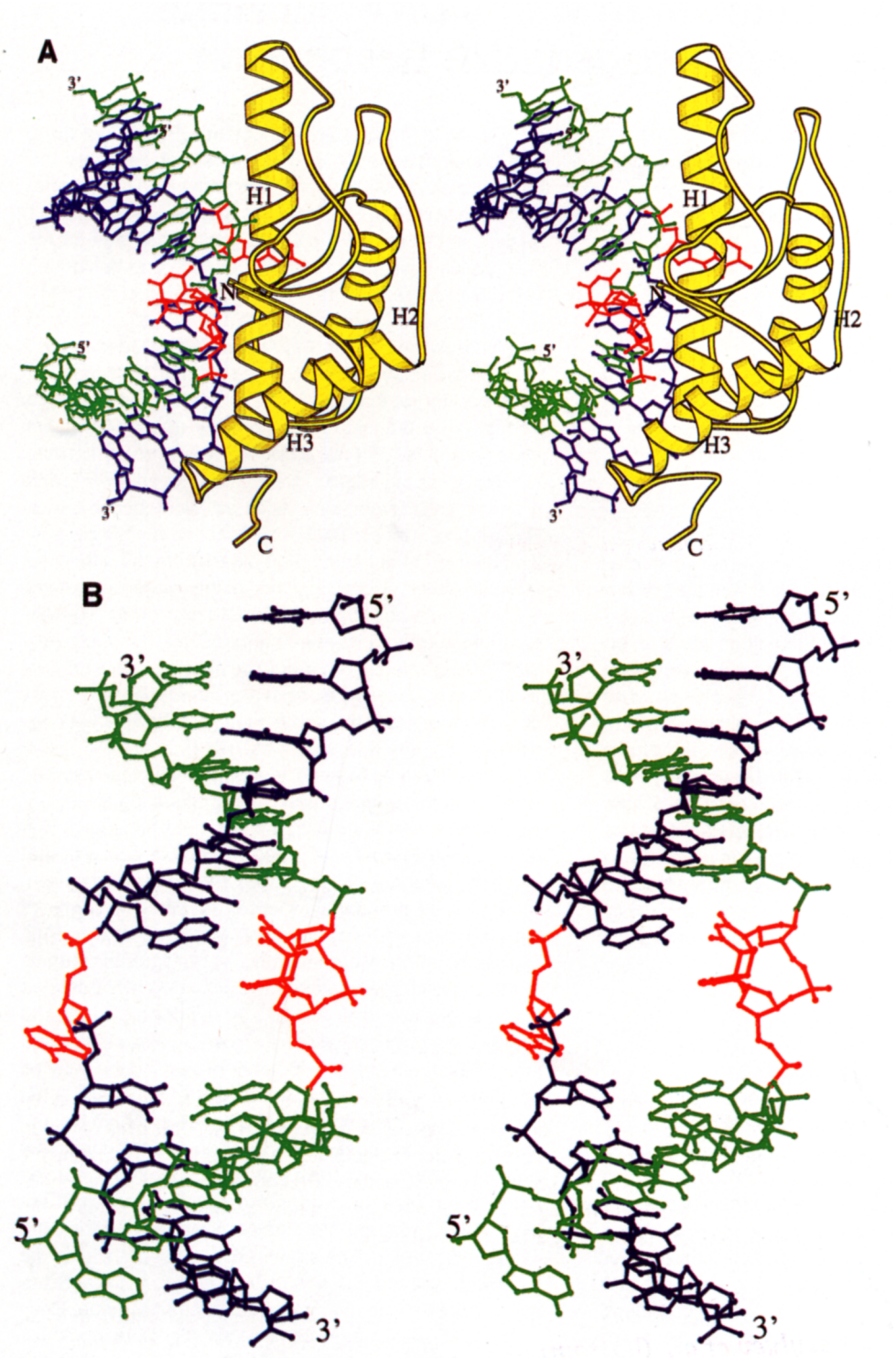


Figure 1. Structure of T4 Endo V–DNA Complex

(A) Stereoview of the T4 endo V–DNA complex drawn with the program MOLSCRIPT (Kraulis, 1991). The protein is represented by a yellow ribbon model. The R and L strands of the DNA substrate are indicated by green and blue ball and stick models, respectively, except for the PD and flipped-out nucleotide, which are colored red. The DNA duplex is viewed perpendicular to its helical axis and toward the sharp kink at the position of the damaged base pair.

(B) Atomic model of the DNA duplex in the T4 endo V–substrate complex (protein part not shown). Coloring is as in (A). The DNA molecule has been rotated by approximately 90° about the longest and vertical axes of the molecule in Figure 3a, such that the kink is coming out at the reader. The figure shows the large hole inside the DNA, which results from the DNA kink and the flipping out of A(L+1). This apparent hole in the complex is actually occupied by protein atoms and some bound water molecules, which are not indicated in this figure.

Table 1. DNA Helical Parameters for DNA Portion Having B-Form Conformation in the Complex with T4 Endo V

		Interbase Pair						Intrabase Pair		
	Base Pair	Tilt (degrees)	Roll (degrees)	Twist (degrees)	Shift (Å)	Slide (Å)	Rise (Å)	Buckle (degrees)	Propeller Twist (degrees)	Open (degrees)
2	T(R–6)–A(L+6)							1.68	–21.12	9.78
3	C(R–5)–G(L+5)	–6.17	7.27	39.15	–0.44	0.02	3.56	–13.63	–6.50	5.15
4	G(R–4)–C(L+4)	–3.33	–0.29	36.80	–0.18	0.30	3.16	2.10	–8.27	3.84
5	C(R–3)–G(L+3)	8.86	0.78	31.68	0.20	–0.24	3.26	12.62	–4.96	3.10
6	G(R–2)–C(L+2)	–6.79	7.57	26.91	0.14	0.35	3.45	30.50	3.86	6.50
8	T(R+1)–A(L–1)	–	–	–	–	–	–	–43.36	1.55	5.77
9	G(R+2)–C(L–2)	2.93	–12.79	34.87	–1.03	0.70	3.19	–2.54	–20.13	4.51
10	C(R+3)–G(L–3)	2.97	–1.27	37.06	0.49	0.41	3.32	–12.92	–11.80	4.41
11	G(R+4)–C(L–4)	–8.87	–0.59	37.29	–0.85	0.97	3.51	–3.85	–16.06	3.42
12	C(R+5)–G(L–5)	1.04	0.05	32.34	–0.20	–0.10	3.30	–1.62	–14.85	2.80
13	T(R+6)–A(L–6)	–1.01	–0.59	34.91	–0.03	0.14	3.56	–3.59	–15.03	5.30

All the parameters listed in the table were calculated by the RNA program (Babcock et al., 1993), which conforms to the standard and common nomenclature for the geometrical DNA parameters. The nucleotides T(R–1) and A(L+1), which do not form a base pair in the complex, were eliminated from the calculations.

Results and Discussion

DNA Conformation

The substrate DNA duplex on the whole adopts a B-DNA structure, which is sharply kinked at the 5′ position of the PD lying at the middle (Figure 1). The helical parameters of the two B-DNA regions within the substrate are summarized in Table 1. The average values of the parameters are, in general, very close to those observed in the DNA dodecamer structure determined by Drew and Dickerson (Drew et al., 1981; Babcock et al., 1993). The average helical twist is 35.4°, which corresponds to 10.2 nt per turn. The twist angles at each step show small deviations, although only the two base pairs C(R–3)–G(L+3) and G(R–2)–C(L+2) (for an explanation of numbering, see the legend of Figure 2) in the close vicinity of the kink exhibit the somewhat unusual helical twist of 26.9°. The average values of the phosphate backbone torsion angles of nucleotides listed in Table 1 are very close to the ideal values for the DNA B-form. Except for the complementary bases, T(R–1) and A(L+1), in the kinked regions, all the nucleotides form the Watson–Crick base pairs. All of the sugar rings exhibit C2′-endo puckering, although slight ambiguities remain in the sugar conformations.

The most remarkable feature of the DNA substrate bound to the enzyme is a sharp kink at the position of the PD, with a 60° inclination made by the two helical axes. This kink appears to be associated with the unusual backbone torsion angles of G(R–2) and the 5′ nucleotide (T[R–1]) of the PD in the R strand and A(L+1) and C(L+2) in the L strand. These unusual backbone conformations result in the destruction of the base pair between T(R–1) and A(L+1). Instead, the A(L+1) base is completely flipped out of the DNA B helix interior. The complementary modified base (T[R–1]) remains within the DNA duplex, although the stacking contact between G(R–2) and T(R–1) is lost.

Another striking feature of the DNA conformation around the photodimer is the apparent formation of a large "hole" with an approximate size of 10 Å × 9 Å (Figure 1B). This unusual architecture is composed of the sugar–phosphate backbones of A(L+1) and T(R–1) and the two base pairs G(R–2)–C(L+2) and T(R+1)–A(L–1). The unusually large values of the buckle angles for these two base pairs (30.5° and –43.3°, respectively; see Table 1) increase the dimensions of the hole (Figure 1B).

The conformation of the thymine dimer coincides well with that of photodimer A of 1,3-dimethylthymine determined by X-ray analysis (Camerman and Camerman, 1970). In contrast with T(R–1), the T(R+1) base makes a strong, but slightly distorted, stacking contact with the neighboring G(R+2). Notably, the spacings between the two phosphates belonging to PD and between the two phosphates of G(R–2) and T(R–1) are 5.2 Å and 5.4 Å, respectively, much smaller than the 6.7 Å found in the normal B-DNA.

In the crystal, the unpaired nucleotides A(R–7) and T(L–7) at the 5′ termini of each strand (Figure 2) form a planar Watson–Crick base pair between two symmetry-related DNA molecules, so that the 12-mer DNA duplexes constitute a continuous superhelix running along a 6_5 screw axis. We believe that this interaction is a major factor stabilizing the molecular arrangements of the complex within the crystal lattice. The formation of continuous helices by DNA oligomers has been found to take place in many other crystals.

Protein Conformation

The conformation of the protein complexed with DNA coincides well with that of the free enzyme (Morikawa et al., 1992, 1994, 1995; Morikawa, 1993). The root mean square displacement values between the two structures are 0.39 Å for all main chain atoms and 0.68 Å for all the atoms including side chains. There are three protein regions in which the binding to DNA induces notable conformational changes.

First, the loop (residues 125–130) near the carboxyl terminus demonstrates substantial differences even in the positions of the main chain atoms. This conformational change appears to be connected with the large move-

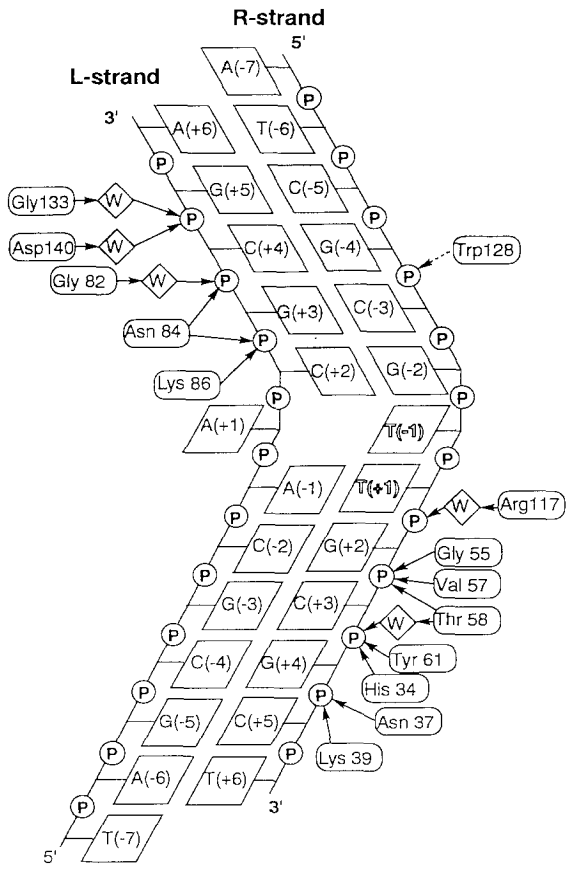


Figure 2. Schematic Drawing of Polar Interactions between T4 Endo V and the DNA Backbone

The thymine dimer is marked by open letters. The DNA strand containing the PD and the complementary strand are designated as the R and L strands, respectively. The numbering of the nucleotides goes in plus (3′) and minus (5′) directions, starting from the position of the PD in the R strand. For instance, the notations represent the 5′ moiety of the PD as T(R−1) and its complementary nucleotide as A(L+1). The stacking contact between Trp-128 and the sugar ring of C(R−3) is indicated by a broken line. The interactions between the protein and the DNA in the R strand are concentrated on the 3′ side of the PD, while the interactions with the L strand are clustered on the 3′ side of the flipped-out base (A[L+1]). This drawing does not include the interactions in the vicinity of the PD, which are represented in Figures 3b and 3c.

ments of the Arg-125 and Trp-128 side chains, which are induced by binding to DNA. This loop partly overlaps with the WYKYY (128–132) segment, which was inferred by mutational analyses to participate in DNA binding (Recinos and Lloyd, 1988; Stump and Lloyd, 1988; Latham et al., 1994).

Significant differences are also observed in the side chain orientation of the residues in the loop 83–91. The structure of this loop is disordered in the free enzyme, as well as in the three mutants (Morikawa et al., 1992, 1994, 1995). In the complex, however, all the residues in this disordered loop are well defined in the map and are involved in DNA binding. The interaction between the disordered segment and the DNA has been inferred on the basis of the previous crystallographic studies (Morikawa et al., 1992, 1994, 1995) and photoaffinity labeling (Hori et al., 1992b).

The third notable structural difference is observed on the Arg-22 and Arg-26 conformations in the catalytic center. In the complex, their side chains approach each other almost at the distance of a hydrogen bond, and both come close to the PD so as to participate directly in the catalytic reaction.

Protein-DNA Interactions

In the overall view of the complex, half of the basic concave surface of the enzyme covers a half region of the B-DNA duplex that is sharply kinked at the central PD, and the remaining half of the enzyme interacts with the other B-DNA portion on the opposite side (Figure 1A). As previously suggested by binding assays using phosphate-modified substrates and DNA methylation analyses of the complex (Iwai et al., 1994), the enzyme interacts with the thymine dimer in the minor groove. The interactions between the enzyme and the DNA substrate are roughly divided into four categories: nonsequence-specific polar interactions with the backbone of the substrate B-DNA; van der Waals interactions of the flipped-out base with the protein atoms in the cavity located on the surface of the enzyme; direct and indirect polar interactions in close vicinity to the PD through extensive hydrogen bond networks between the basic side chains and the deformed phosphate backbones; and polar interactions of protein atoms with the PD bases and sugar moieties in the active site. Upon binding of the enzyme to the substrate, these interactions would take place simultaneously and in a concerted manner.

Interactions with B-DNA Regions

There are ten direct and five water-mediated polar interactions between the protein and the two B-DNA regions (Figure 2).

On the 3′ side of the the PD, the enzyme enters into nine polar interactions with the phosphate backbone of the R strand. These extensive polar interactions induce no distortion of the B-DNA conformation, but restrain the orientation of the local helical axis of the DNA duplex on the 3′ side of the R strand. Notably, the enzyme has no interaction with the L strand at all, ranging from the 5′-terminal nucleotide (T[L−7]) to the flipped-out base (A[L−1]). Six similar polar interactions occur in the other B-DNA half on the 5′ side of the PD. However, the interactions involve only the backbone atoms of the L strand of the DNA, and none of the atoms of the R strand is involved.

Thus, on each side of the B-DNA duplexes, either the R (G[R+2]–T[R+6]) or the L strand (G[L+3]–A[L+6]) is anchored to the protein. The orientation of each strand is dominated by the extensive polar interactions with the concave surface of the enzyme, as mentioned above. These anchors would additionally restrain the positions of the respective strands on the interface of the protein. Then, the positions of the respective complementary strands, T(L−7)–A(L−1) and A(R−7)–C(R−3), are spontaneously fixed by the formation of Watson–Crick base pairs to con-

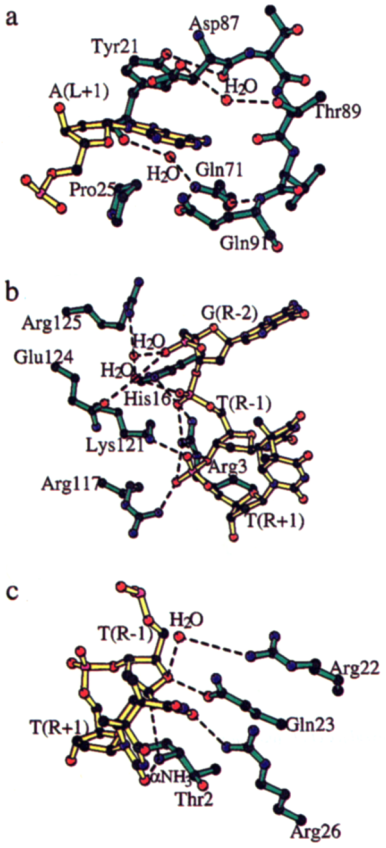


Figure 3. Protein–DNA Interactions in the Vicinity of the PD

The atomic model is represented by balls and sticks. Balls have atom-dependent colors, and sticks are colored green and yellow for the protein and nucleotide moieties, respectively. The drawing was produced by MOLSCRIPT (Kraulis, 1991).

(a) Environment around the flipped-out base (A[L+1]). The protein atoms making van der Waals contacts with A(L+1) are shown. The hydrogen bond networks, which form two layers to sandwich the adenine base, are shown by broken lines. One layer, 3.5 Å away from the base plane, is composed of the phenol ring of Tyr-21, the Cβ, Cγ, Oδ1, and Oδ2 atoms of Asp-87, a water molecule (H_2O-432), and the Oγ1 of Thr-89. The second layer, which makes contact with the opposite side of the base approximately 3.3 Å distant from the base plane, consists of several atoms, such as the O of Tyr-21, the Cγ of Pro-25, a water molecule (H_2O-402), the N, Cδ, Oε1, and Nε2 atoms of Gln-91, and the Cγ, Cδ, Oε1, and Nε2 atoms of Gln-71.

(b) Polar interactions of the enzyme with the phosphate backbone near the PD. The amino acid residues taking part in the PD-specific recognition are shown. The polar interactions, between the protein and the respective phosphates belonging to the PD and between the protein and the phosphates adjacent to the 5′ side of the PD, are shown by broken lines.

(c) Active site of T4 endo V (E23Q mutant) in the complex with DNA. The important amino acid residues are shown. The possible contacts (not hydrogen bonds) between the respective protein atoms and the PD are shown by broken lines. A polar interaction of the α-amino group of Thr-2 with the CO_2 of the 3′ thymine ring (T[R+1]) of the PD would fix its active position. The Nη2 atom of Arg-22 indirectly interacts with the O4′ atom of the T(R−1) deoxyribose, through a mediating water molecule (H_2O-396). The side chain of Arg-26 forms a hydrogen bond with the CO_2 of T(R−1). In the inactive complex, the substituted amide group of Gln-23 makes three polar contacts with the T(R−1) residue: one with the CO_2 of the base and two with O4′ of the sugar ring, and hence the carboxyl side chain of Glu-23 would lie in close vicinity to the T(R−1) base and sugar. The crystallographic study combined with mutational analyses (Doi et al., 1992; Morikawa et al., 1994, 1995; Manuel et al., 1995) revealed that the carboxyl side chain of Glu-23 plays a crucial role in the cleavage of the N-glycosyl bond.

struct the normal B-DNA. Thus, the strikingly distorted architecture in the close vicinity of the PD is able to coexist with the normal B-DNA in the complex.

Trapping of the Flipped-Out Adenine Base

In the complex, the base of A(L+1) complementary to T(R−1) (5′ moiety of the PD) is completely flipped out from the interior of the DNA duplex and is accommodated into the cavity, with approximate dimensions of 10 Å × 7 Å × 9 Å, on the protein surface (Figure 3a). Interestingly, in the cavity the adenine base does not form any hydrogen bonds with the protein atoms, although the polar side chains of Asp-87, Gln-71, and Gln-91 approach the adenine ring. Instead, this base is sandwiched between two layers of protein atoms, both of which are arranged in parallel with the base plane. We assume that the major force for trapping the base would be weak van der Waals interactions with the surrounding atoms.

These structural features indicate that the interactions within the cavity are nonspecific and that, in principle, any other aromatic bases could be trapped in a similar manner. Consistent with this hypothesis, recent analyses of cleavage efficiencies of modified DNA substrates (M. Maeda, unpublished data) have demonstrated that the replacement of A(L+1) by G, C, or T does not affect the glycosylase activity. However, it should have a critical role in preventing the slippage of the DNA duplex along the helical axis. Furthermore, the generation of the DNA kink would be coupled with the flipping out of the complementary base to the 5′ side of the PD, and hence the major role of the cavity may be the maintenance of a sufficient space to accommodate the flipped-out base.

Specific Recognition within the Vicinity of the PD

One might expect that the enzyme would recognize the unique architecture produced by the dimer formation, such as the cross-linked moiety of the pyrimidine rings or the violation of the base planarity (Figures 3b and 3c), as was proposed for the photolyase (Park et al., 1995). Contrary to this expectation, we find no interaction at all between these atoms and the enzyme. On the other hand, there are extensive interactions between the amino acid side chains and the deformed phosphate backbones in the vicinity of the PD. The most striking feature of the deformed backbone is the spacing between the two phosphates of the PD, which is shortened by 1.5 Å as compared with the normal B-DNA. Notably, these two phosphates enter into seven direct polar interactions with Arg-3, His-16, Arg-117, and Lys-121 (Figure 3b). The side chain of Arg-3 forms four hydrogen bonds with the two phosphates and probably plays the key role in the PD recognition, which is in good agreement with the mutational analysis of the enzyme (Doi et al., 1992). In the complex, Gln-23 in the active site does not contribute to the substrate recog-

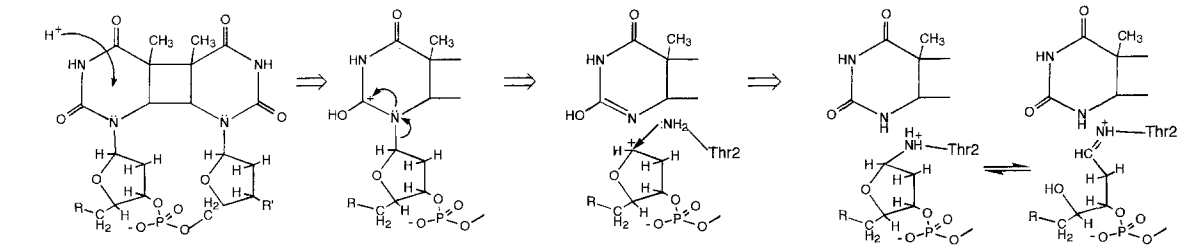
nition. This may explain why the inactive E23Q mutant retains the same affinity to the substrate as the wild-type enzyme (Doi et al., 1992).

Similarly, the distance between the two adjacent phosphates belonging to G(R−2) and T(R−1) becomes substantially shorter. These phosphates also serve in the specific recognition of the substrate by the enzyme. However, in contrast with the above case, this recognition is accomplished mainly through the "indirect readout," which involves water-mediated hydrogen bonds (Figure 3b).

It should be noted that Arg-26, among the active site residues, contributes to the PD recognition. The side chain of this residue forms stacking contact with Ade(L−1), which probably stabilizes the unusual buckle angle of the Ade(L−1)–Y(R+1) base pair (Table 1) and thus defines the correct orientation of the PD backbone to the protein.

Active Site and Implications for the Catalytic Mechanism

Biochemical studies have revealed that an imino–covalent enzyme–substrate intermediate is formed between the terminal α-amino group of Thr-2 and the C1′ of the 5′ deoxyribose of the PD, and this major finding has led to a proposal for the catalytic scheme (Schrock and Lloyd, 1991; Dodson et al., 1993; Iwai et al., 1995). The structural features of the complex are generally consistent with this scheme, although we raise the possibility of slight modifications so as to reconcile it with the various interactions found in this study (Figures 3c and 4). We can implicate four residues, Arg-22, Glu-23, Arg-26, and the amino-terminal Thr-2, as participants in the catalytic reactions. Consistent with this implication, site-directed mutagenesis experiments demonstrated that mutations of Glu-23 completely abolish the enzymatic activity, while the replacements of Arg-22 and Arg-26 cause the striking decrease of the activity (Doi et al., 1992). The side chains of Arg-22, Gln-23 (Glu-23 in the wild type) and Arg-26, occupy the empty space within the DNA duplex that is generated by the flipping out of Ade(L+1). Although Arg-22 and Arg-26 are involved in interactions with the 5′ moiety of PD (Figure 3c), it is unlikely that they are responsible for donating a proton to either the pyrimidine ring or the sugar of the target nucleotide (Figure 4) because of the high pK_a values of guanidino groups. These two residues probably play an important structural role during the cleavage reaction of the glycosyl bond in securing the correct positioning of the thymine ring and the deoxyribose of T(R−1). The terminal α-amino group lies at a 3.8 Å distance from the C1′ atom of the 5′ deoxyribose of the PD (Figure 3c), which makes it capable of acting as an attacking nucleophile in the initial catalytic step. The positively charged Schiff base intermediate can be stabilized by the side chain of Glu-23 (Schrock and Lloyd, 1991; Manuel et al., 1995). This carboxyl side chain may have an alternative role in the protonation of the pyrimidine ring, thereby causing the N-glycosyl bond to be labile. These two possible functions of Glu-23 may explain why the E23Q mutant is completely inactive. Furthermore, this residue appears to participate in β elimination. In agreement with the findings by mutational analysis (Hori et al., 1992a; Manuel et al., 1995), a tentative model (data not shown) of the Schiff base intermediate implies that the pro-S2 hydrogen at the C2′ position of the open deoxyribose is abstracted by the negatively charged carboxyl side chain of Glu-23. A similar model-building approach also



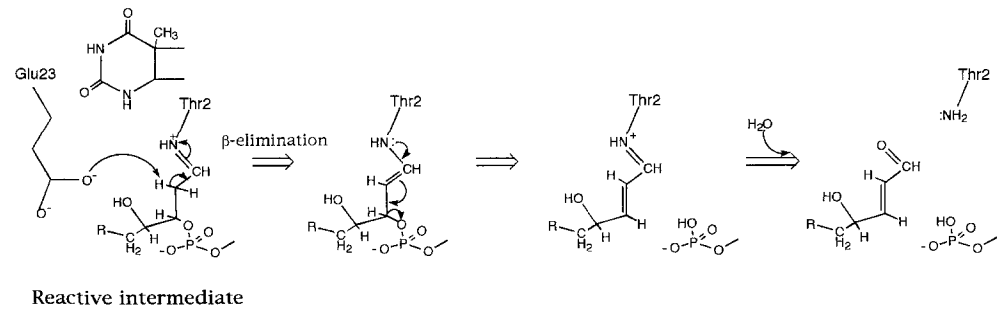


Figure 4. Catalytic Mechanism of T4 Endo V

This reaction scheme largely relies upon the proposal by Schrock and Lloyd (1991). The possible dissociation pathway of this reaction, which occurs at high pH, remains the same as their scheme and hence is not shown here.

indicates that the final product possesses a small spacing (3.2 Å) between the negatively charged phosphate oxygens adjacent to the α,β-unsaturated aldehyde and the 5′ monophosphate. The repulsive force derived from this spacing could break the hydrogen bond networks (Figure 3b), which are a major factor connecting the enzyme and DNA.

Recognition of the DNA Lesion

We should stress three major structural factors of the damaged DNA that are recognized by the enzyme. The first factor is the sharp kink at the central PD portion, which splits the DNA duplex into two halves of B-DNA. We assume that this remarkable kink would be induced by the binding of the DNA substrate to the enzyme, since NMR analyses of a DNA oligomer containing a thymine photodimer demonstrated that the normal B-DNA double helical structure, as a whole, remains intact (Kemmink et al., 1987a, 1987b; Taylor et al., 1990; Lee et al., 1994). Nonspecific electrostatic interaction between the protein and the DNA backbone (Dodson et al., 1994) may induce this kink. Presumably, the weakened stacking interaction at the PD moiety (Kemmink et al., 1987a, 1987b; Lee et al., 1994) would facilitate this deformation. Consistent with this assumption, it has recently been reported that T4 endo V can bind a DNA duplex containing an apyrimidinic site in a manner similar to that for the PD-containing DNA (Latham et al., 1995). However, we cannot exclude the possibility that the kink, present as a minor conformational variant of the DNA lesion before binding, is initially recognized by the enzyme and all the DNA are then driven into the conformation observed in the complex. Although a smooth bend in the DNA has been observed in many protein–DNA complexes, a sharp kink so far has been found only in four other cases: the catabolite gene activator protein–DNA complex (Schultz et al., 1991); the EcoRV endonuclease–cognate DNA complex (Winkler et al., 1993); the TATA-binding protein–DNA complex (Kim et al., 1993a; Kim et al., 1993b); and, more recently, the purine repressor–DNA complex (Schumacher et al., 1994). However, all of these cases exhibit a smaller value of inclination (about 45°) than the 60° of the T4 endo V–DNA complex, and they never exhibit the remarkable disruption of the base pair seen in the present case.

The second is the flipping out of the adenine base complementary to the 5′ PD moiety. This flipping out plays two important roles in the reaction catalyzed by T4 endo V: on one side of the strand, the enzyme can discriminate the damaged DNA from normal DNA duplexes through the trapping of the flipped-out base into the protein cavity. On the other side, the flipping out generates an empty space inside the DNA, which is occupied by catalytically active residues in the complex. A similar flipped-out base was also observed in the methyltransferase (MT)–DNA complex (Cheng et al., 1993; Klimasaukas et al., 1994). There are three major differences between these two complex structures. In the case of the MT–DNA complex, a loop of the enzyme undergoes a large conformational change upon binding of the substrate, whereas the T4 endo V–DNA substrate complex shows no substantial alteration of the enzymic structure. The former complex retains a straight and normal B-DNA conformation and does not indicate a remarkable deformation like the sharp kink in the T4 endo V–DNA substrate. Furthermore, in the PD recognition, only the complementary adenine base is flipped out, leaving the damaged moiety inside the DNA duplex. In contrast, the flipped-out cytosine base in the MT is the direct target for the catalytic reaction. In spite of the large difference in the DNA conformations between T4 endo V–DNA and the MT–DNA complexes, the flipping out in both cases enables the active amino acid residues to gain access to the target nucleotide. Most recently, a similar possibility of a flipped-out base has been proposed for the uracil-DNA glycosylases (Sarva et al., 1995; Mol et al., 1995). Thus, the flipped-out nucleotide may be a general feature of DNA deformation induced by enzymes that act to modify DNA bases.

Third, T4 endo V recognizes the deformation of the phosphate backbone in the vicinity of PD induced by the DNA kink, which is typically represented by the two kinds of shortened spacings between the adjacent phosphates. In fact, these phosphates have extensive polar interactions with the enzyme (Figure 3b).

Some findings indicate that another DNA-repair enzyme, endo III from E. coli, may bind to DNA and recognize its substrates in a manner similar to that of T4 endo V, although these two enzymes have no similarity in their primary structures, main chain foldings, or substrate specificities. For instance, NMR studies (Kao et al., 1993) showed that, even in the absence of the enzyme, the substrate DNA for endo III has a significant distortion of the B-DNA at the position of the damaged base, which could serve for the initial recognition of damaged DNA by the enzyme. The crystal structure of the DNA-free enzyme (Kuo et al., 1992) reveals that the loop 112–121, which was proposed to contain catalytically active residues, lies relatively far from the basic surface of the enzyme. This surface should be involved in the nonspecific binding of the DNA duplex. Thus, it is likely that the damaged nucleotide is extrahelical to gain access to the catalytic center of the enzyme. Furthermore, the substrate specificity of endo III has been found to be very broad (Kow and Wallace, 1987). Since the enzyme is unlikely to recognize so many different types of damaged bases, it presumably would recognize the DNA substrate through the unique phosphate backbone deformation, which in turn could be induced by the DNA kink.

This mechanism, in which the damage itself within the DNA is not recognized by the protein, tempts us to assume that it may be common to other DNA repair enzymes.

Experimental Procedures

Preparation of DNA Substrate

The oligonucleotide containing the thymine dimer was prepared by the phosphoramidite method, using a dinucleotide-coupling unit, as described previously (Taylor et al., 1987) and was purified by reverse phase high performance liquid chromatography. The DNA duplex

ATCGCGT$\widehat{T}$GCGCT
AGCGCAACGCGAT

Table 2. Summary of Crystallographic Analysis

Data Collection

Device	Resolution (Å)	Number of Reflections: Total	Number of Reflections: Unique	R_{sym} (%)
DIP100	50.0–2.75	24,907	5,662	6.2

Molecular Replacement

Rotation function

	θ_1(degrees)	θ_2(degrees)	θ_3(degrees)	PC Correlation Function
First peak	124.20	159.70	0.58	0.077
Second peak	306.80	20.74	55.20	0.042

Translation function

	Space Group	Position (Fraction of Unit Cell): X	Y	Z	Translation Function
First peak	$P6_5$	0.451	0.089	0.000	0.426
Second peak	$P6_5$	0.785	0.755	0.000	0.220
First peak	$P6_1$	0.451	0.084	0.000	0.178

Refinements Statistics

Resolution (Å)	10–2.75	5.26	4.28	3.77	3.44	3.20	3.02	2.87	2.75
Complete (%)	71.0	87.3	89.9	87.0	81.3	72.4	61.4	47.8	39.0
Reflections	5447	872	866	844	778	691	569	465	362
R factor (%)[b]	15.2	17.7	14.1	13.3	13.4	14.8	16.0	17.8	20.1

Number of nonhydrogen atoms	1799
Rms bond lengths (Å)	0.015
Rms bond angles (degrees)	1.98
Rms improper angles (degrees)	1.20

[a] $R_{sym} = \Sigma\Sigma|I - \langle I\rangle|/\Sigma\Sigma\langle|$, where I is observed intensity and $\langle I\rangle$ is average intensity, obtained from multiple observations of symmetry-related reflections.

[b] R factor = $\Sigma|Fo - Fc|/\Sigma Fo$, where Fo and Fc are observed and subsequently calculated structure factors, respectively. Rms, root mean square.

was produced by mixing the thymine dimer–containing oligonucleotide with the complementary strand at a molar ratio of 1:1.1 and subjecting them to an annealing procedure.

Crystallization of the T4 Endo V-DNA Complex

An active site mutant, E23Q (Doi et al., 1992; Morikawa, 1993; Morikawa et al., 1994, 1995), which retains almost the same K_m value with a DNA substrate as the wild-type enzyme but which lacks catalytic activity (Doi et al., 1992), was used for the cocrystallization. The previous crystallographic study confirmed that this replacement produced little conformational change, even in the active site (Morikawa et al., 1994, 1995). The protein was purified as described previously (Doi et al., 1992). The crystal of the complex with a 13-mer DNA duplex, which contains a thymine dimer and a 5′ unpaired base in each strand, was grown using the hanging-drop vapor diffusion method at 4°C. The crystallization drop (10 μl) containing 0.38 mM protein, 0.53 mM DNA duplex, 150 mM Tris–HCl (pH 7.0), and 33% saturated $(NH_4)_2SO_4$ was equilibrated with a 1 ml reservoir solution (54% saturated $[NH_4]_2SO_4$). Hexagonal prismatic crystals appeared in about 2 months.

Data Collection and Processing

X-ray data to 2.75 Å resolution were collected at room temperature with a DIP100 imaging plate diffractometer (MAC Science) equipped on a rotating anode X-ray generator operating at 45 kV × 90 mA. After data collection, the space group and unit cell parameters were determined from the same crystal, using a CAD4 four circle diffractometer (Enraf Nonius). The crystal belonged to the $P6_1$ (or $P6_5$) space group with the following unit cell parameters: a = b = 118.82 Å; c = 36.28 Å. It contains one complex molecule per an asymmetric unit, and the solvent content is roughly 50%. The statistics of data collection are summarized in Table 2.

Structure Determination

The molecular replacement (MR) technique, using the X-PLOR program package (Brunger, 1993), was applied to the data to obtain the correct orientation and position of the protein molecule in the complex crystal. The coordinates of the wild-type enzyme structure refined at 1.45 Å resolution (Morikawa et al., 1994, 1995) were used for the vector search calculation. The translation function was jointly used to solve the ambiguity of the space group. Both rotation and translation searches yielded single strong peaks for the $P6_5$ space group (Table 2). The rigid body refinement was then carried out, and, subsequently, a |2Fo–Fc| electron density map was calculated. This map showed clear densities for the protein molecule, but the connectivities of the electron densities in the DNA region were so poor that the trace for the DNA molecule could not be achieved. Two algorithms for electron density modification were used to improve the MR phases: solvent flattening (SF), using programs from the CCP4 suit (SERC Collaborative Computing Project Number 4, a suit of programs for protein crystallography; Daresbury Laboratory, Warrington, England), and skeletonization (SK), which uses the program PRISM (Baker et al., 1993) and auxiliary programs from CCP4. When starting with the MR phases, the SF did not improve phases effectively, presumably owing to inaccurate determination for the molecular envelope of the DNA. In contrast, the SK procedure allowed substantial improvement of the electron densities in the DNA region without any information regarding the solvent envelope. Thus, the correct envelope, corresponding to a 35% solvent content, was obtained, and the SF and SK procedures were repeated with this new envelope. At this point, both SF and SK provided similarly improved electron densities, which allowed unambiguous model building of the entire DNA molecule. The average discrepancy of the phases was 33° between MR and SK and 39° between MR and SF. During the SK calculations, the value of the nodes (>1.2σ of electron density)

in the skeleton rose from 1100 to 8800, and the percentage of the nodes in the largest graph was increased from 13% to 57%. It is worthwhile to mention that a particular procedure of molecular masking for a known portion in the electron density map, which is followed by phase addition in reciprocal space, was performed throughout all the calculations. The molecular mask was calculated from the coordinates of a protein model, using the atom spheres of 2.5 Å. At each cycle of density modification, the electron density inside the molecular mask was fixed to zero, whereas the density outside the mask was modified. At the same time, the structure factors calculated from the protein model were kept constant, and in each cycle they were added in the reciprocal space by structure factors, which an electron density map modified by SF or SK yields. This procedure could impose strong restraints against phase deviation from the initial values and led to a great improvement at the undefined regions in the electron density map without disturbing the densities in the known part of the map. The model building of the DNA was carried out using the O program (Jones et al., 1991). The nucleotides were carefully fitted into electron densities one by one, because the map still showed several breaks in the densities corresponding to the sugar moieties or the phosphate backbones and also because DNA were so substantially kinked that we could not directly fit a straight B-DNA model into them. The densities for the thymine dimer were initially fitted with the crystal structure solved by X-ray analysis (Camerman and Camerman, 1970), and, subsequently, it was used to form the standard chemical groups in the topology and parameter files of X-PLOR. All the refinements were carried out according to the simulated annealing protocol in X-PLOR. After the first round of X-PLOR refinement, the R factor decreased from 37% to 20%, and the corresponding electron densities were greatly improved. However, the geometry of the DNA was somewhat unsatisfactory because of mistakes made during the manual model building and the different orders of the force constants in the X-PLOR parameter files between amino acid and nucleotide residues. The force constants in the DNA parameter file were scaled to that for the protein, and the DNA model was revised manually using the program O to obtain conventional values for the backbone torsion angles and the hydrogen bonds between the bases. Using the new model, the refinement was continued under the strong restraints for DNA backbone torsion angles, hydrogen bond distances for base pairs, and also sugar puckering. Owing to these restraints, all the parameters specified above were settled to almost ideal values, except for the base pairs in the position of the PD. At the final stage of refinement, the water molecules found in the |Fo–Fc| electron density map were included, and the restraints were omitted from the calculations. The water molecules with B factor values greater than 50 were removed from the final model. The current model includes 137 amino acids, 26 base DNA constructs, and 143 water molecules. The average B factor values for the protein molecule and the DNA molecule are 7.8 Å and 31 Å, respectively. Both the protein and the DNA are clearly represented by electron densities in the final |2Fo–Fc| map, in spite of the relatively large average temperature factor for the DNA. In the complex structure, only T(R–6) shows slightly ambiguous electron densities with respect to the sugar–phosphate backbone. The final R factor is 15.2% (R free = 24.8%) for all data with Fo > 2σ(F), and the atomic model has excellent stereo chemistry (Table 2).

Acknowledgments

Correspondence should be addressed to K. M. We are grateful to Tetsuya Murata, Tatsuya Osafune, and Mie Maeda for their help in the synthesis of the DNA substrate. Dr. Satoshi Fujii is acknowledged for his valuable suggestions of DNA structures, and Prof. Morio Ikehara is thanked for his support.

Received August 17, 1995; revised October 12, 1995.

References

Babcock, M.S., Pednault, E.P.D., and Olson, W.K.J. (1993). Nucleic acid structure analysis: a user guide to a collection of new analysis programs. Biomol. Struct. Dyn. *2*, 597–628.

Baker, D., Bystroff, C., Fletterick, R.J., and Agard, D.A. (1993). PRISM: topologically constrained phase refinement for macromolecular crystallography. Acta Cryst. *D49*, 429–439.

Brunger, A.T. (1993). X-PLOR Version 3.1 (New Haven, Connecticut: Yale University Press).

Camerman, N., and Camerman, A. (1970). Crystal and molecular structure of a photodimer A of 1,3-dimethylthymine (the isomer in irradiated deoxyribonucleic acid). J. Am. Chem. Soc. *92*, 2523–2527.

Cheng, X., Kumar, S., Posfai, J., Pflugrath, J.W., and Roberts, R.J. (1993). Crystal structure of the HhaI DNA methyltransferase complexed with S-adenosyl-L-methionine. Cell *74*, 299–307.

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

Dodson, M.L., Michaels, M.L., and Lloyd, R.S. (1994). Unified catalytic mechanism for DNA glycosylases. J. Biol. Chem. *269*, 32709–32711.

Doi, T., Recktenwald, A., Karaki, Y., Kikuchi, M., Morikawa, K., Ikehara, M., Inaoka, T., Hori, N., and Ohtsuka, E. (1992). Role of the basic amino acid cluster and Glu-23 in pyrimidine dimer glycosylase activity of T4 endonuclease V. Proc. Natl. Acad. Sci. USA *89*, 9420–9424.

Drew, H.R., Wing, R.M., Takano, T., Broka, C., Tanaka, S., Itakura, K., and Dickerson, R.E. (1981). Structure of a B-DNA dodecamer: conformation and dynamics. Proc. Natl. Acad. Sci. USA *78*, 2179–2183.

Hamilton, K.K., Kim, P.M.H., and Doesch, P.W. (1992). A eukaryotic DNA glycosylase/lyase recognizing ultraviolet light-induced pyrimidine dimers. Nature *356*, 725–728.

Hori, N., Doi, T., Karaki, Y., Kikuchi, M., Ikehara, M., and Ohtsuka, E. (1992a). Participation of glutamic acid 23 of T4 endonuclease V in the β-elimination reaction of an abasic site in a synthetic duplex DNA. Nucl. Acids Res. *20*, 4761–4764.

Hori, N., Iwai, S., Inoue, H., and Ohtsuka E. (1992b). Photoaffinity labeling of T4 endonuclease V with a substrate containing a phenyldiazirine derivative. J. Biol. Chem. *267*, 15591–15594.

Inaoka, T., Ishida, M., and Ohtsuka, E. (1989). Affinity of single- or double-stranded oligodeoxyribonucleotides containing a thymine photodimer for T4 endonuclease V. J. Biol. Chem. *264*, 2609–2614.

Iwai, S., Maeda, M., Shimada, Y., Hori, N., Murata, T., Morioka, H., and Ohtsuka, E. (1994). Endonuclease V from bacteriophage T4 interacts with its substrate in the minor groove. Biochemistry *33*, 5581–5588.

Iwai, S., Maeda, M., Shirai, M., Shimada, Y., Osafune, T., Murata, T., and Ohtsuka, E. (1995). Reaction mechanism of T4 endonuclease V determined by analysis using modified oligonucleotide duplexes. Biochemistry *34*, 4601–4609.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Cryst. *A47*, 110–119.

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

Kemmink, J., Boelens, R., Koning, T.M.G., Kaptein, R., van der Marel, G.A., and van Boom, J.H. (1987a). Conformational changes in the oligonucleotide duplex d(GCGTTGCG)·d(CGCAACGC) induced by formation of a *cis-syn* thymine dimer. Eur. J. Biochem. *162*, 37–43.

Kemmink, J., Boelens, R., Koning, T., van der Marel, G.A., van Boom, J.H., and Kaptein, R. (1987b). ^{1}H NMR study of the exchangeable protons of the duplex d(GCGTTGCG)·d(CGCAACGC) containing a thymine photodimer. Nucl. Acids Res. *15*, 4645–4653.

Kim, J., and Linn, S. (1988). The mechanism of action of *E. coli* endonuclease III and T4 UV endonuclease (endonuclease V) at AP sites. Nucl. Acids Res. *16*, 1135–1141.

Kim, J.L, Nikolov, D.B., and Burley, S.K. (1993a). Co-crystal structure of TBP recognizing the minor groove of a TATA element. Nature *365*, 520–527.

Kim, Y., Geiger, J.H., Hahn, S., and Sigler, P.B. (1993b). Crystal structure of a yeast TBP/TATA-box complex. Nature *365*, 512–520.

Klimasaukas, S., Kumar, S., Roberts, R.J., and Cheng, X. (1994). HhaI methyltransferase flips its target base out of DNA helix. Cell *76*, 357–369.

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

Kraulis, P.J. (1991). MOLSRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Cryst. *24*, 946–950.

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

Latham, K.A., and Lloyd, R.S. (1994). T4 endonuclease V: perspectives on catalysis. In DNA Damage, S. Wallace, B.V. Houten, and Y.W. Kow, eds. (New York: New York Academy of Sciences), pp. 181–197.

Latham, K.A., Carmical, R., and Lloyd, R.S. (1994). Mutation of tryptophan 128 in T4 endonuclease V does not affect glycosylase or abasic site lyase activity. Biochemistry *33*, 9024–9031.

Latham, K.A., Manuel, R.C., and Lloyd, R.S. (1995). The interaction of T4 endonuclease V E23Q mutant with thymine dimer- and tetrahydrofuran-containing DNA. J. Bacteriol. *177*, 5166–5168.

Lee, B.J., Sakashita, H., Ohkubo, T., Ikehara, M., Doi, T., Morikawa, K., Kyogoku, Y., Osafune, T., Iwai, S., and Ohtsuka, E. (1994). Nuclear magnetic resonance study of the interaction of T4 endonuclease V with DNA. Biochemistry *33*, 57–64.

Macmillan S., Edenberg, H.J., Radany, E.H., Friedberg, R.C., and Friedberg, E.C. (1981). *denV* gene of bacteriophage T4 codes for both pyrimidine dimer-DNA glycosylase and apyrimidinic endonuclease activities. J. Virol. *40*, 211–223.

Manoharan, M., Mazumder, A., Ranson, S.C., Gerlt, J.A., and Bolton, P.H. (1988). Mechanism of UV endonuclease V cleavage of abasic sites in DNA determined by ^{13}C labeling. J. Am. Chem. Soc. *11*, 2690–2691.

Manuel, R.C., Latham, K.A., Dodson, M.L., and Lloyd, R.S. (1995). Involvement of glutamic acid 23 in the catalytic mechanism of T4 endonuclease V. J. Biol. Chem. *270*, 2652–2661.

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

Morikawa, K. (1993). DNA repair enzymes. Curr. Opin. Struct. Biol. *3*, 17–24.

Morikawa, K., Matsumoto, O., Tsujimoto, M., Katayanagi, K., Ariyoshi, M., Doi, T., Ikehara, M., Inaoka, T., and 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., Katayanagi, K., Nakamura, H., Doi, T., Hori, N., and Ohtsuka, E. (1994). Crystal structure of T4 endonuclease V: an excision repair enzyme for a pyrimidine dimer. In DNA Damage, S. Wallace, B.V. Houten, and Y.W. Kow, eds. (New York: New York Academy of Sciences), pp. 198–207.

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

Nakabeppu, Y., and Sekiguchi, M. (1981). Physical association of pyrimidine dimer DNA glycosylase and apurinic/apyrimidinic DNA glycosylase essential for repair of ultraviolet-damaged DNA. Proc. Natl. Acad. Sci. USA *78*, 2742–2746.

Park, H.-W., Kim, S.-T., Sancar, A., and Deisenhofer, J. (1995). Crystal structure of DNA photolyase from *Escherichia coli*. Science *268*, 1866–1872.

Recinos, A., and Lloyd, R.S. (1988). Site directed mutagenesis of the T4 endonuclease V gene: role of lysine-130. Biochemistry *27*, 1832–1838.

Sancer, A., and Sancer, G.B. (1988). DNA repair enzymes. Annu. Rev. Biochem. *57*, 29–67.

Sarva, R., Katherline, M.-H., Brown, T., and Pearl, L. (1995). The structural basis of specific base-excision repair by uracil-DNA glycosylase. Nature *373*, 487–493.

Schrock, R.D., III, and Lloyd, R.S. (1991). Reductive methylation of the amino terminus of endonuclease V eradicates catalytic activities. J. Biol. Chem. *266*, 17631–17639.

Schrock, R.D., III, and Lloyd, R.S. (1993). Site-directed mutagenesis of the NH_2 terminus of T4 endonuclease V. J. Biol. Chem. *268*, 880–886.

Schultz, S.C., Shields, G.C., and Steitz, T.A. (1991). Crystal structure of CAP-DNA complex: the DNA is bent by 90°. Science *253*, 1001–1007.

Schumacher, M.A., Choi, K.Y., Zalkin, H., and Brennan, R.G. (1994). Crystal structure of LacL member, PurR, bound to DNA: minor groove binding by α-helixes. Science *266*, 763–770.

Stump, D.G., and Lloyd, R.S. (1988). Site-directed mutagenesis of the T4 endonuclease V gene: role of tyrosine-129 and -131 in pyrimidine dimer-specific binding. Biochemistry *27*, 1839–1848.

Taylor, J.S., Brockie, I.R., and O'Day, C.L. (1987). A building block for the sequence-specific introduction of cis-syn thymine dimers into oligonucleotides: solid-phase synthesis of TpT[c,s]pTpT. J. Am. Chem. Soc. *109*, 6735–6742.

Taylor, J.S., Garret, D.S., Brockie, I.R., Svoboda, D.L., and Tesler, J. (1990). ^1H NMR assignment and melting temperature study of *cis-syn* and *trans-syn* thymine dimer containing duplexes of d(GCTATTATGC)·d(GCATAATACG). Biochemistry *29*, 8858–8866.

Winkler, F.K., Banner, D.W., Oefner, C., Tsernoglou, D., Brown, R.S., Heathman, S.P., Bryan, R.R, Martin, P.D., Petratos, K., and Wilson, K.S. (1993). The crystal structure of EcoRV endonuclease and of its complexes with cognate and non-cognate DNA fragments. EMBO J. *12*, 1781–1795.