A Model for the Enzyme−Substrate Complex of DNA Photolyase and Photodamaged DNA

Don B. Sanders and Olaf Wiest*

Contribution from the Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556-5670

Received July 27, 1998

Abstract: The three-dimensional structure of *Escherichia coli* DNA photolyase and molecular dynamics simulations using the AMBER force field were used to construct a model of the enzyme−substrate complex. Three different dinucleotides with cyclobutane pyrimidine dimers (**T**<>**T**, **T**<>**U**, and **U**<>**T**), two conformations of a single-stranded DNA nonamer, and a duplex DNA dodecamer containing the **T**<>**T** lesion were studied. The results are in good agreement with available experimental data and provide a structural rationalization for the results of ethylation studies, the measurement of the relative rates of electron transfer for different dinucleotides complexed to the enzyme, and the similar binding constants for **T**<>**T** containing single stranded and duplex DNA. The results support the base-flipping mechanism suggested earlier. The proposed active-site model reveals three types of interactions: (i) ion-pair interactions at the rim of the active site between the positively charged residues on the enzyme surface (Arg^{226}, Arg^{342}, Arg^{397}, and Lys^{154}) and the deoxyribophosphate immediately 5′ to the dimer as well as the three deoxyribophosphates on the 3′ side, (ii) polar interactions between Glu^{274} and the NH function of the 3′ base of the dimer as well as a hydrogen bond between the C-4 carbonyl on the 5′ base of the dimer with Trp^{384}, and (iii) hydrophobic interactions between Trp^{277} and Trp^{384} and the nonpolar cyclobutane moiety of the dimer, thus shielding the radical anion intermediate of the DNA repair from electrophilic attack. In this model, the distance between the redox active FADH cofactor and the dimer is too large to account for the observed rates of electron transfer. Rather, the results suggest an electron transfer mediated by the π-systems of the aromatic residues Trp^{277} and Trp^{384}.

Introduction

The principal damage in DNA caused by ultraviolet radiation is the formation of cyclobutane pyrimidine dimers (CPD). Irradiation of two adjacent thymine bases **T** in the DNA with light of wavelengths between 260 and 320 nm induces a [2+2] cycloaddition to form the *cis,syn* cyclobutane thymine dimer **T**<>**T**[1] (Figure 1). These covalently linked dimers block cell replication and transcription, therefore causing cell death. It is also widely accepted that the presence of **T**<>**T**[2] and another of the minor photoproducts, the 6−4 photoadduct,[3] are major causes for basal and squamous cell skin cancers in humans. Skin cancer is now the most common form of cancer in the US with an estimated 1.4 million diagnosed cases in the US alone.[4]

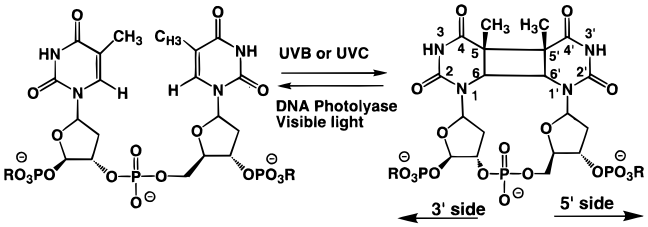

Figure 1. Formation and cycloreversion of the *cis,syn* thymine dimer.

The number of patients has been steadily rising over the last 20 years. A further increase is to be expected due to the rapid depletion of the earth's protective ozone layer. It is therefore not surprising that the formation and the further fate of **T**<>**T** has been the topic of intense studies over the last five decades.[1,5,6]

The photoinduced damage of DNA can be repaired through a light-induced repair mechanism, mediated by the enzyme DNA photolyase. Contrary to the replacement of damaged DNA fragments by the more common excision enzymes, these

* Author to whom correspondence may be addressed. E-mail: wiest.1@nd.edu.

(1) For recent monographs, see, for example: (a) Friedberg, E. C. *DNA Repair*; W. H. Freeman & Co.: New York, 1985; Chapters 1−5, 2−1, and 2−2. (b) Cadet, J.; Vigny, P. The Photochemistry of Nucleic Acids. In *Bioorganic Photochemistry of Nucleic Acids*; Wang, S. Y., Ed.; Academic Press: New York, 1976; Vol. I, pp 226−295. (c) Patrick, M. H.; Rahn, R. O. *Photochemistry and Photobiology of Nucleic Acids*; Wang, S. Y., Ed.; Academic Press: New York, 1976; Vol. II, pp 35−95.

(2) See, for example: (a) Alcalay, J.; Freeman, S.E.; Goldberg, L. H.; Wolf, J. E. *J. Invest. Dermatol.* **1990**, *95*, 506−509. (b) Hart, R. W.; Setlow, R. B. In *Molecular Mechanisms for Repair of DNA*; Hannawalt, P. C., Setlow, R. B., Eds., Plenum: New York, 1975; p 719. (c) Setlow, R. B. *Photochem. Photobiol.* **1997**, *65*, 119−122. (d) Taylor, J.-S. *Acc. Chem. Res.* **1994**, *27*, 76−82.

(3) (a) Brash, D. E.; Haseltine, W. A. *Nature* **1982**, *298*, 189−192. (b) Mitchell, D. L. *Photochem. Photobiol.* **1988**, *48*, 51−57.

(4) (a) Goldberg, L. H. *Lancet* **1996**, *347*, 663−667. (b) Marks, R. *Lancet* **1996**, *347*, 735−738. (c) Taylor, S.-J. *Pure Appl. Chem.* **1995**, *67*, 183−190.

(5) See, for example: (a) Friedberg, E. C. *Trends Biochem Sci* **1995**, *20*, 381. (b) Koshland, D. E. *Science* **1994**, *266*, 1925. (c) Strauss, B. S. *Nature* **1993**, *366*, 301. (d) Hearst, J. E. *Science* **1995**, *268*, 1858−1859. (e) Carell, T. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 2491−2494.

(6) For recent overviews compare, for example: (a) Sancar, A.; Sancar, G. B. *Annu. Rev. Biochem.* **1988**, *52*, 29−67. (b) Sancar, A. *Biochemistry* **1994**, *33*, 2−9. (c) Sancar, A. In *Advances in Electron-Transfer Chemistry*; Mariano, P. S., Ed.; JAI Press: New York, 1992; Vol. 2, pp 215−272. (d) Begley, T. B. *Acc. Chem. Res.* **1994**, *27*, 394−401. (e) Heelis, P. F.; Hartman, R. F.; Rose, S. D. *Chem. Soc. Rev.* **1995**, *24*, 289−297.

10.1021/ja982660y CCC: $18.00
Published on Web 05/22/1999

remarkable enzymes achieve a true repair through an electron-transfer catalyzed [2 + 2] cycloreversion of **T<>T**.[6] Photolyases have been found in many organisms from all three biological kingdoms, but their presence in humans is still a matter of debate.[7,8] They are monomeric proteins with a molecular weight between 55 and 65 kD and contain two noncovalently bound cofactors. One of them, a 1,5-dihydroflavin adenine dinucleotide ($FADH_2$) common to all known photolyases, is positioned in the active site of the enzyme. The completely reduced form, $FADH^-$, acts as the electron donor in the electron-transfer step. The other cofactor, which is either methenyltetrahydrofolate (MTHF) or 8-hydroxy-5-deazaflavine (8-HDF), depending on the organism from which the DNA photolyase was isolated, acts as a light-harvesting antenna and transfers the light energy through a Förster-type energy transfer to the $FADH_2$ cofactor. Photolyases bind single- or double-stranded DNA containing **T<>T** with binding constants of 10^{-8} M and 10^{-9} M, respectively. Other photoadducts are bound less tightly, but the dinucleotide of **T<>T** is bound with a K_D of 10^{-14} M.[9] Substrate binding is essentially independent of the DNA sequence.[10,11]

The structure of DNA photolyase from *Escherichia coli*, a member of the MTHF class, has recently been determined by X-ray crystallography.[12] The overall structure consists of five parallel β-strands, 20 α helices, and five short 3_{10} helices which are organized in three major domains. One domain is formed by the β-strands and four α helices that bind the second cofactor, MTHF. The other two, formed by the helical structures, are separated by a hole in the otherwise relatively flat protein surface. This hole, containing the noncovalentely bound flavine cofactor, approximates the dimensions of a CPD and constitutes the putative binding site. The rim of the binding site is positively charged (Arg^{226}, Arg^{342} and Arg^{397}), whereas the interior is uncharged and consists of relatively hydrophobic amino acid residues (Phe^{150}, Val^{270}, Trp^{277}, Tyr^{281}, Met^{345}, Trp^{384}, and Ala^{392}). The hole lies in the middle of a series of solvent-exposed, positively charged amino acid residues designed to bind to the phosphate backbone of DNA.

The three-dimensional structure of the enzyme−substrate complex of DNA photolyase is not available. This is unfortunate since this information could provide insights into some key questions such as the mode of substrate recognition, the electron transfer from the FADH cofactor to the **T<>T** lesion, and the question of how the enzyme copes with the highly reactive radical anion intermediate. Knowledge of the relative orientation of the substrate in the active site at a molecular level could increase our knowledge of this unique DNA repair enzyme. The wealth of information available from the X-ray structure and from binding studies can be exploited with the use of molecular modeling. DNA photolyase is particularly suitable for such a docking study because the position of the active site is known and the pronounced electrostatic pattern of the solvent-exposed surface around the active site limits the number of possible relative orientations of the enzyme and the substrate. Here, we report a model of the three-dimensional structure of the enzyme−substrate complex that is in agreement with the available experimental data. After an outline of the methods used for model building and refinement, we will discuss the binding of several different cyclobutane pyrimidine dimers to the active site. We will then present our results for the binding of both a single strand and duplex DNA fragment containing a **T<>T** to DNA photolyase with and without base flipping. Finally, we will discuss the implications for other DNA photolyases.

Computational Methods

Model Building and Refinement. All calculations were performed using the AMBER force field[13] as implemented in Biosym's DISCOVER3, with a nonbonded atom-based cutoff of 9.5 Å and a distance-dependent dielectric constant of 1.0. A combination of steepest descent and Polak−Ribiere conjugate gradient minimizers with a 0.001 kcal/mol convergence criterion was used in all final minimizations. All MD simulations were run in a canonical ensemble using a Verlet velocity integrator with a 1-fs time step.

The crystal structure of DNA photolyase from *E. coli* was obtained from the Brookhaven Protein Database (structure code 1DNP). Hydrogens were added, and the potentials of the cofactors were set to match known atoms within the AMBER force field (see Supporting Information). The DNA photolyase model was then surrounded by a 5-Å layer of water and refined using restrained minimization, followed by a 3-ps MD simulation and unrestrained minimization. As a model for the photodamaged DNA, we used a dinucleotide containing **T<>T** and the adjacent deoxyribophosphates as well as a single strand DNA nonamer CGAAT<>TCGC. **T<>T** was introduced by changing the atom types at the C5/C5′ and C6/C6′ positions to "CT" atom type of AMBER and adjusting the parameters for bond lengths, angles, and torsions according to the values reported in previous theoretical studies.[14,15] All nonstandard parameters used are listed in the Supporting Information. The validity of these parameters was tested by constructing the previously studied (CGCGAAT<>TCGCG) (CGCGAATTCGCG) DNA duplex in the B-form with sodium counterions in a 61 Å x 61 Å x 61 Å waterbox with periodic boundary conditions. The surrounding water was equilibrated and the initially used constrains on the Watson−Crick pairing in the duplex were gradually removed as described earlier[16] until a completely minimized structure was obtained. After 70 ps of MD simulations, followed by another minimization, we obtained a structure with a 15−20° kink in the DNA strand. This is in good agreement with the results from considerably longer MD studies by Spector et al.,[16] who obtained a 22.3° kink. Another recent study of the same dodecamer by Miaskiewicz gave an overall kink of ∼15°.[17] It can therefore be concluded that, for the system studied here, there are only small differences between the results obtained by the versions of the AMBER force field used here and in earlier studies.[16,17] Experimentally, early gel electrophoresis studies indicated a bend of approximately 30°, while newer electrophoresis and NMR experiments found a smaller kink of 7−9°.[18]

(7) (a) Li, Y. F.; Kim, S.-T.; Sancar, A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 4389−4393. (b) Ley, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 4337.

(8) (a) Sutherland, B. M.; Bennet, P. V. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9732−9736. (b) deGruijl, F. R.; Roza, L. *J. Photochem. Photobiol. B: Biol.* **1991**, *10*, 367−371.

(9) (a) Kim, S.-T.; Sancar, A. *Photochem Photobiol.* **1993**, *57*, 895−904. (b) Sancar, G. B.; Smith, F. W.; Sancar, A. *Biochemistry* **1985**, *24*, 1849−1855. (c) Husain, I.; Sancar, A. *Nucleic Acids Res.* **1987**, *15*, 1109−1120.

(10) Myles, G. M.; VanHouten, B.; Sancar, A. *Nucleic Acids Res.* **1987**, *15*, 1227−1243.

(11) Payne, G.; Sancar, A. *Biochemistry* **1990**, *29*, 7715−7727.

(12) (a) Park, H.-W.; Kim, S.-T.; Sancar, A.; Deisenhofer, J. *Science* **1995**, *268*, 1866−1872. (b) Park, H.-W.; Sancar, A.; Deisenhofer, J. *J. Mol. Biol.* **1993**, *231*, 1122−1123.

(13) Weiner, S. J.; Kollman, P. A.; Case, D.; Singh, U. C.; Ghio, C.; Alague, G., Profecta, S., Jr.; Weiner, P. K. *J. Am. Chem. Soc.* **1984**, *106*, 765−784.

(14) (a) Rao, S. U.; Keepers, J. W.; Kollman, P. A. *Nucleic Acids Res.* **1984**, *12*, 4789−4807. (b) Kim, S.-H.; Pearlman, D. A.; Holbrook, S. R.; Pirkle, D. In *Molecular Basics of Cancer, Part A: Macromolecular Structure, Carcinogens, and Oncogenes*; Liss: New York, 1985; pp 143−152.

(15) Wiest, O.; Largoza, A.; *Internet J. Sci.−Biol. Chem.* **1997**, 1.

(16) Spector, T. I.; Cheatham, T. E.; Kollman, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 7095−7104.

(17) Miaskiewicz, K.; Miller, J.; Cooney, M.; Osman, R. *J. Am. Chem. Soc.* **1996**, *118*, 9156−9163.

(18) (a) Husain, I.; Griffith, J.; Sancar, A. *Proc. Natl. Acad. Sci.U.S.A.* **1988**, *85*, 2558−2562. (b) Wang, C.-I.; Taylor, J.-S.. *Proc. Natl. Acad. Sci.U.S.A.* **1991**, *88*, 9072−9076. (c) Kim, J.-K.; Patel, D.; Choi, B.-S. *Photochem. Photobiol.* **1995**, *62*, 44−50. (d) McAteer, K.; Jing, Y.; Kao, J.; Taylor, J.-S.; Kennedy, M. A. *J. Mol. Biol.* **1998**, *282*, 1013−1032.

Docking Procedures. Initial docking studies were performed on a truncated model of the enzyme using residues 209−408 and the FADH cofactor. This approximately corresponds to the residues within 10 Å of the active site. Connoly surfaces of the enzyme and the DNA piece with approximate electrostatic potentials projected on the surfaces were generated by the DELPHI module of the Biosym software. By using these surfaces and the electrostatic and van der Waals interaction grids generated by the DOCKING module of Biosym, the DNA piece was manually docked with the enzyme active site. Approximately 30 different relative orientations were considered in the manual docking. Once a local minimum was found, the truncated enzyme was replaced by the full enzyme, and a minimization using all AMBER interaction terms was performed. The backbone of the enzyme was fixed, and only amino acids near the putative active site (Arg^{226}, Arg^{342}, Arg^{397}, Val^{270}, Trp^{277}, Tyr^{281}, Met^{345}, Trp^{384}, and the $FADH_2$ cofactor)[12a] were allowed to move freely. This procedure was followed both for CGAAT<>TCGC nonamer, and for a **T**<>**T** with adjacent deoxyribophosphate groups. The best relative dinucleotide−enzyme and DNA−enzyme complexes were chosen for a considerably longer dynamics procedure. This longer procedure consisted of sequence of unconstrained minimization of the complete enzyme, a 10-ps MD run at 298 K, followed by 80 ps of equilibration at 350 K.[19] After this preparatory phase, data were collected every 5 ps for 120 ps at 350 K. The relatively high temperature was chosen to enable the system to overcome barriers toward lower energy structures. The simulation was then finished by another 10-ps MD simulation at 298 K and unconstrained minimization of all atoms in the low energy structure obtained in the MD simulations.

Although the relatively short MD simulation time of 200 ps is too short to allow for a comprehensive exploration of the conformational space, the pronounced electrostatic pattern around the active site and the fact that we arrived at the same structures from approximately 30 quite different starting structures indicates that the number of substantially different local minima is fairly limited. In addition, we will show in this paper that the proposed model is in good agreement with the available experimental data. We therefore conclude that the model is at the very least semiquantitatively correct and provides the structural basis for the interpretation of the mode of binding.

Results and Discussion

The Enzyme−Dinucleotide Model. We used the known three-dimensional structure of DNA photolyase from *E. coli* to construct a model of the enzyme−substrate complex. This structure had been refined to a resolution of 2.3 Å ($R = 0.172$) with the position of only one out of 471 amino acids undefined. After the addition of the hydrogen atoms and water molecules, this structure was minimized as described above. Comparison of the atom positions in the fully minimized structure with the original X-ray data gave a rms displacement of 1.004 Å for the entire enzyme. Most of the differences occurred by the movement of side chains of amino acids near the surface, since these have the greatest amount of freedom. Very little movement occurred to the backbone itself.

We started our investigations by the manual docking of a dinucleotide containing **T**<>**T** into the active site. The position of the active site is well-established by the position of the FADH cofactor. Furthermore, Park et al.[12a] pointed out that the phosphate backbone of the damaged DNA strand most likely runs parallel to a streak of positive charge which runs along the surface of the photolyase over the active site. These previous studies allowed us to limit the manual docking studies to approximately 30 orientations for the dinucleotide−photolyase interactions for a reasonable exploration of the possible binding modes. These starting structures were chosen to explore numerous different relative orientations of the dinucleotide in the active site. By refining the starting structures as described above, all structures converged independently of the starting structure to one of two low-energy conformations. These were subjected to full energy minimization without constraints, followed by MD simulations to yield the two final structures. Figure 2 shows a detailed view of the active site of the lower-energy model.

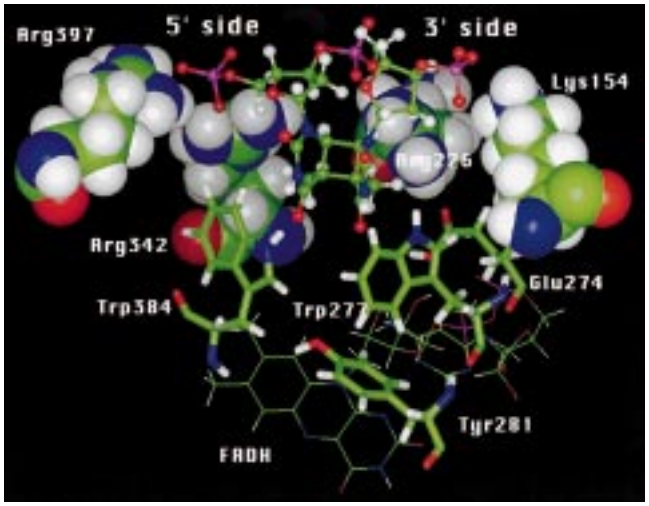

Figure 2. Active site of the low-energy enzyme−substrate complex model. The dinucleotide containing **T**<>**T** is shown in ball-and-stick representation, charged amino acid residues on the surface of the enzyme as CPK models, nonpolar amino acid residues are shown in tube representation, and the FADH cofactor is shown as a wiredrawing.

Two types of interactions between the dinucleotide and the enzyme become apparent in this structure. The deoxyribophosphate backbone on the 5′ side of the **T**<>**T** lesion interacts strongly with Arg^{342} and Arg^{397}, which almost trap the phosphate group, effectively restricting any degrees of freedom. The phosphate on the 3′ side has similarly strong interactions with Lys^{154}. With the adjacent phosphate groups essentially anchored by these two interactions, **T**<>**T** can only rotate along the axis between the two arginines and Lys^{154}. These strong interactions are consistent with the findings that alkylation of either the 3′ or the 5′ phosphates immediately adjacent to **T**<>**T** decreases binding.[20] In addition, we also found ion-pair interactions between Arg^{226} and the phosphate group linking the two thymines. Since experimental evidence suggests that this phosphate group is not important for binding, we assume that it is well-solvated in the actual enzyme−substrate complex. Due to its position outside the active site, we do not anticipate a major influence of this interaction on the orientation of **T**<>**T** in the active site.

In addition to these mostly ion-pair interaction outside and at the rim of the active site, we also obtained several nonpolar and dipolar interactions within the active site. Our model shows a polar contact between the side chain of Glu^{274} and the N3 of the thymine on the 3′ side as well as a possible hydrogen bond between Trp^{384} and the carbonyl function at C4 of the 5′ base of the **T**<>**T**. This hydrogen bond could stabilize the localized charge in the radical anion, thus lowering the redox potential of **T**<>**T**. Similar effects of hydrogen bonding on reduction potentials have been observed for a number of small model systems.[21] The **T**<>**T** in the active site is in contact with the two tryptophans, Trp^{277} and Trp^{384}, which create a fairly hydrophobic binding pocket for the CPD. This hydrophobic

(19) This equilibration time is similar to the one chosen in another recent MD study of *E. coli* photolyase: Hahn, J.; Michel-Beyerle, M.-E.; Rösch, N. *J. Mol. Model* **1998**, *4*, 73−82.

(20) (a) Husain, I.; Sancar, G.; Holbrook, S.; Sancar, A. *J. Biol. Chem.* **1987**, *262*, 13188−13197. Compare also: (b) Liuzzi, M.; Weinfield, M.; Paterson, M. *J. Biol. Chem.* **1989**, *264*, 6355−6363.

pocket is presumably required not only to interact with the hydrophobic cyclobutane moiety of the molecule but also to shield the highly reactive radical anion intermediates from electrophilic attack.

It is interesting to note that there are no close contacts between the **T**<>**T** and the FADH cofactor. The smallest distance, between the C5 on the 3′ base of the **T**<>**T** and the ribose side chain, is larger than 6 Å, whereas the smallest distance between **T**<>**T** and the redox active isoalloxazine ring system is around 10 Å. This distance changes very little (±0.3 Å) during the 120 ps of data collection.[22] Several attempts to locate a structure with a smaller distance have not been successful, since such a structure would require the disruption of the strong ion-pair interactions on the surface of the enzyme. Reoptimization of any of these structures therefore led back to the original geometry. A direct electron transfer over this relatively large distance should therefore be much slower than the observed overall rate constant for repair of the **T**<>**T** lesion of 5×10^9.[23] Instead, our model supports an earlier proposal by Kim et al.,[24] according to which the electron transfer is mediated by the π-systems of Tyr^{281} and Trp^{277} or by Trp^{384}. Both pathways are possible in our enzyme−substrate complex model. Such a mediation of electron transfer by the π-systems of aromatic amino acid residues is analogous to that by a number of other redox active biomolecules most notably the photosynthetic reaction center. The experimental evidence for the involvement of these residues is, however, inconsistent. Site-directed mutagenesis of Trp^{277} in *E. coli* photolyase changes the binding constant of the enzyme but does not have a significant effect on electron transfer.[25] In contrast, mutagenesis of Trp^{387}, the equivalent residue in yeast photolyase, reduced the quantum yield by a factor of 3.[26] Recently, a model qualitatively similar to the one proposed here was obtained by docking a pyrimidine dimer to a model of the experimentally not yet determined active-site structure of yeast photolyase.[27] Unfortunately, no site-directed mutagenesis results are available for Trp^{384} in *E. coli* photolyase.

In addition to the enzyme−dinucleotide model discussed above, a second low-energy structure could be located through the MD−minimization protocol. This structure results from a flip of the direction of the DNA. When the dimer is reversed in the 3′−5′ direction, similar ion-pair interactions between the deoxyribophosphate backbone and the positively charged amino acid residues on the surface are obtained, but other interactions in the active site are not as favorable. In particular, the contacts between the CPD and Trp^{277} and Trp^{384} are disrupted since the CPD is rotated by almost 180° in the active site, resulting in steric repulsion of the tryptophans and the methyl groups of **T**<>**T**. Therefore, and because this orientation of the **T**<>**T** does not reproduce the experimental results of the alkylation studies (also see below),[20] this model was not further pursued.

In a recent study of four different CPDs, Langenbacher et al. showed that DNA photolyase activity depends on the substrate.[28]

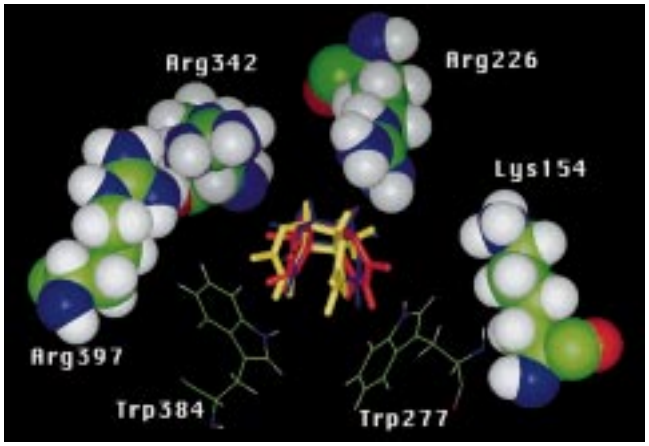

Figure 3. Overlay of the relative positions of the **U**<>**T** (gold), **T**<>**T** (blue), and **T**<>**U** dimers (red) in the active-site model. The deoxyribophosphate backbone has been removed for clarity.

By using picosecond spectroscopy, they found that the rate of electron transfer to either **U**<>**U** or **U**<>**T** is faster than the rate of electron transfer to either **T**<>**U** or **T**<>**T**. They suggested that the position of **T**<>**U** in the enzyme−substrate complex is less favorable for electron transfer than the one of **U**<>**T**. The methyl group in the 5′ base of the **T**<>**T** would alter the position of the base, thus decreasing the rate of electron transfer. The importance of the position of the 5′ base indicates, in accordance with earlier findings,[29] that the electron in the radical anion is localized in the 5′ rather than in the 3′ base.

To study the applicability of our model to these findings, we calculated the enzyme−substrate complex models for the **T**<>**U** and **U**<>**T** dinucleotides. An overlay of the results is shown in Figure 3. Removal of the methyl group in the 5′ base results in a clockwise rotation of the **U**<>**T** dinucleotide, shown in gold, thus bringing the 5′ base closer to Trp^{384}. In the **T**<>**U** system, shown in red, the steric repulsion of the methyl group with Arg^{342} backbone leads to a 5−10° counterclockwise rotation of the CPD. The **T**<>**T** dinucleotide, shown in blue, lies approximately between the two. It is interesting to note that the smaller distance between the 5′ pyrimidine and Trp^{384} correlates with the experimentally observed higher rate of electron transfer. This could indicate that Trp^{384} is more important than Trp^{277} in the alternative pathway. The relative positions of the CPD as a function of substitution at C-5 is therefore in good agreement with the observed behavior, indicating that the different rate constants of electron transfer are indeed due to changes in the electronic coupling matrix element **V** between the FADH cofactor and **T**<>**T**, as suggested earlier.[28]

The Enzyme−Single Strand DNA Model. To study the influence of the DNA backbone on the structure of our model of the enzyme−substrate complex, we investigated the interaction of a nine-nucleotide strand of DNA (CGAAT<>TCGC) with DNA photolyase. There have been several studies showing that DNA photolyase binds and repairs single- and double-stranded DNA equally well.[9,30] It is widely accepted that the mode of binding for single- and double-stranded DNA is very similar and that there are few contacts between the enzyme and the DNA strand opposite the **T**<>**T** lesion.[6b,9,12a,31] Single-

(21) (a) Breinlinger, E. W.; Niemz, A.; Rotello, V. M. *J. Am. Chem. Soc.* **1995**, *117*, 5379−5380. (b) Ge, Y.; Lilienthal, R. R., Smith, D. K. *J. Am. Chem. Soc.* **1996**, *118*, 3976−3977.

(22) Measured from N1 of the isoalloxazine ring to C5 of **T**<>**T**.

(23) Heelis, P. F., Okamura, T.; Sancar, A. *Biochemistry* **1990**, *29*, 5694−5698 and ref 6c.

(24) (a) Kim, S.-T.; Heelis, P. F.; Sancar, A. *Methods Enzymol.* **1995**, *258* 319−343. For a related model of a putative electron transfer from Trp^{306} to FADH, compare: Cheung, M. S.; Daizadeh, I.; Stuchebrukhov, A. A.; Helis, P. F. *Biophys. J.* **1999**, *76*, 1241−1249.

(25) Li, Y. W.; Sancar, A. *Biochemistry* **1990**, *29*, 5698−5706.

(26) Baer, M. E.; Sancar, G. B. *J. Biol. Chem.* **1993**, *268*, 16717−16724.

(27) Vande Berg, B. J.; Sancar, G. B. *J. Biol. Chem.* **1998**, *273*, 20276−20284.

(28) Langenbacher, T.; Zhao, X.; Bieser, G.; Heelis, P.; Sancar, A.; Michel-Beyerle, M. *J. Am. Chem. Soc.* **1997**, *119*, 10532−10536.

(29) Pouwels, P. J.; Hartman, R. F.; Rose, S.; Kaptein, R. *Photochem. Photobiol.* **1995**, *61*, 575−583.

(30) Jorns, M. S.; Sancar, G. B.; Sancar, A. *Biochemistry* **1985**, *24*, 1856−1861.

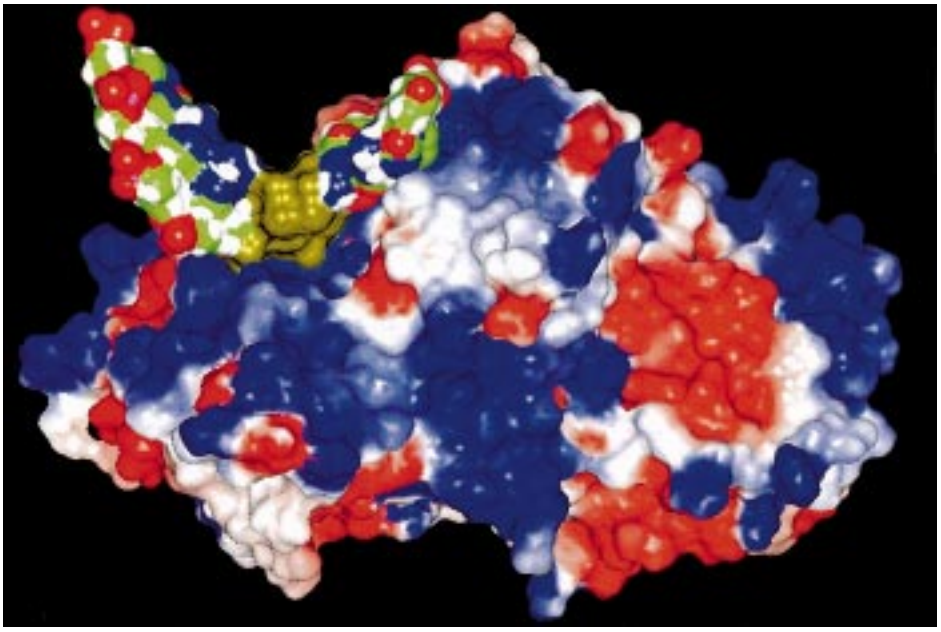

Figure 4. Complex of DNA photolyase with the approximate electrostatic potential projected onto the Connolly surface, complexed to a nine-base single strand DNA piece. The **T**<>**T** fragment is shown in gold.

stranded DNA should therefore be a good model for the biologically relevant DNA duplex. We began our studies with a single-stranded DNA piece containing **T**<>**T** in the same position it occupies in a double helix. Since in this orientation **T**<>**T** is pointing away from the enzyme, the interactions between the photolyase active site and the thymine dimer itself were poor. Following a protocol of manual docking, minimization, molecular dynamics, and reminimization, we obtained the structure shown in Figure 4. In this representation, the **T**<>**T** lesion is highlighted in gold. The 3′ side of the nonamer is positioned right of the lesion, the 5′ side is to the left.

It is noteworthy that, although obtained independently from the results of the interactions in dinucleotide−enzyme system, the interactions between the deoxyribophosphate backbone of the nine-nucleotide strand of DNA and the solvent-exposed surface of the enzyme are essentially identical to the interactions discussed earlier. As in the dinucleotide−enzyme system, strong interactions occurred between the phosphate groups on either side of the dimer with Arg^{397} and Arg^{342} and Lys^{154}. More favorable overall interactions were obtained when the phosphate group immediately 5′ to the **T**<>**T** lesion was paired with Arg^{397} and Arg^{342} than when the 3′ phosphate group was. These are exactly the same results we obtained in the dinucleotide−enzyme system, similar to the interactions shown in Figure 2. In analogy to the systems discussed earlier, further interactions occur between Arg^{226} and the phosphate group immediately 3′ to the **T**<>**T** lesion and between Arg^{342} and the phosphodiester group connecting the two thymines.

Such surface interactions in DNA repair enzymes are not unprecedented. T4 endonuclease V is a DNA repair enzyme from bacteriophage T4 which catalyzes the first step of a thymine dimer-specific base excision. Vassylyev et al. determined the crystal structure of the enzyme−substrate complex and analyzed the role of surface contacts for binding and repair of **T**<>**T**.[32] This system has interactions similar to those which occur in our model of the enzyme−substrate. Most interactions occur on the 3′ side of the lesion and include an arginine interacting with the phosphodiester bond immediately 3′ to the dimer. Other interactions occur with the second, third, and fourth phosphodiester groups 3′ to the dimer and between a tryptophan and the second phosphodiester group 5′ to the dimer.

Husain et al. reported that ethylating any one of the three phosphodiester bonds on the 3′ side of the **T**<>**T** lesion reduces the binding.[20] In contrast, only ethylation of the phosphate bond immediately 5′ to **T**<>**T** interferes with binding. The origin of these findings are readily apparent from our model. It can be seen from Figure 4 that strong surface interactions occur in our system between the second and third phosphodiester bonds 3′ to **T**<>**T** and the enzyme. The distances from the phosphate groups on either side of **T**<>**T** to the surface-exposed residues of the enzyme are both less than 2 Å. The distances of the second and third deoxyribophosphates on the 3′ side of **T**<>**T** to the enzyme surface residues are approximately 3 and 6 Å, respectively. Consequently, ethylation of any of the three phosphates on the 3′ side will interfere with binding by disruption of the ion-pair interactions on the surface. In comparison, only the phosphodiester immediately to the 5′ side of **T**<>**T** has significant surface contacts in our model, whereas the remaining part of the DNA strand has no surface contacts. Ethylation of these phosphate bonds is therefore expected to have only a small effect on binding.

In this model, the dimer is situated over the active site but makes few contacts in the active site and is shifted in the 5′ direction when compared with the dimer−enzyme system. The phosphate group of the DNA strand immediately 5′ to the dimer lies directly over Arg^{397}, while in the dinucleotide−enzyme system, the 5′ phosphate group was located further down into the active site and is in contact with Arg^{397} (compare Figure 7). This shift is further demonstrated by the interactions that occur between Arg^{226} and the phosphate group 3′ to the dimer. This same amino acid residue interacts with the phosphate group linking the two thymines in the dinucleotide−enzyme system. As a consequence of this shift, the contacts of the dimer to Trp^{277} and Trp^{384} are disrupted, and the dimer is even further removed from the redox active center. This shift may in part be the result

(31) Sancar, G. B. *Mutat. Res.* **1990**, *236*, 147−160.
(32) Vassylyev, D.; Kashiwagi, T.; Mikami, Y.; Ariyoshi, M.; Iwai, S.; Ohtsuka, E.; Morikawa, K. *Cell* **1995**, *83*, 773−782.

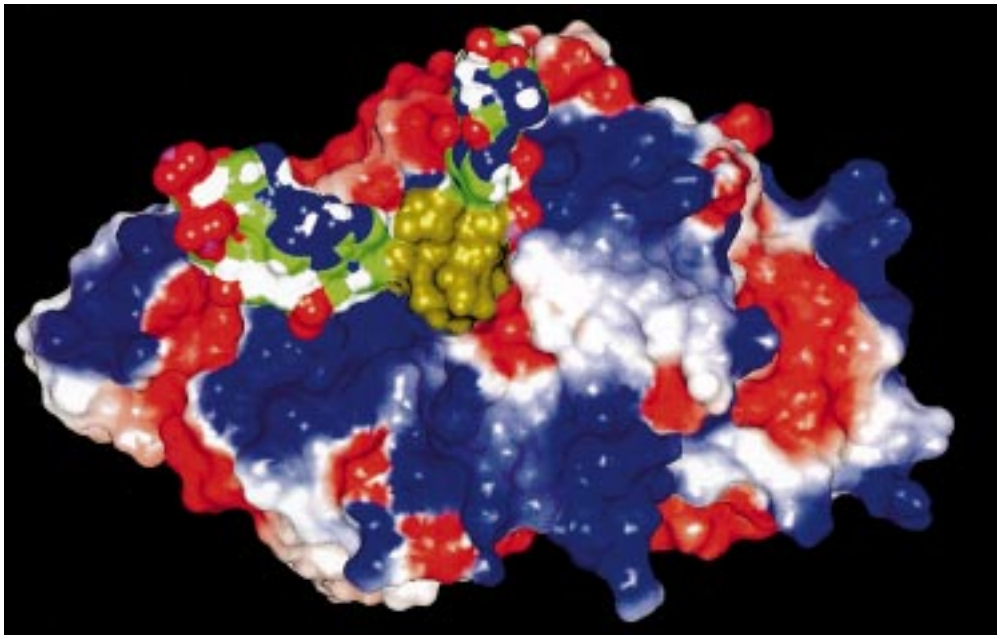

Figure 5. Complex of DNA photolyase with the approximate electrostatic potential projected onto the Connolly surface, complexed to a nine base single-stranded DNA piece. The **T**<>**T** fragment, shown in gold, has been flipped out of the original position.

of the additional interactions on the 3′ side of the DNA single strand which are not present in the dinucleotide−photolyase model.

The poor dimer−active-site interactions obtained by this model give further support to the proposal by Park et al. that **T**<>**T** flips out of the DNA strand upon binding to the enzyme.[12a] Such a base flip has been observed in a number of DNA−protein complexes,[33] including the complex of T4 endonuclease and thymine dimer containing DNA,[32] and has been shown to occur in the case of the yeast photolyase.[27] This flip is easily possible because of the locally disrupted Watson−Crick base pairing in the damaged DNA. In the endonuclease T4 complex, the base opposite the lesion moves out of the normal double strand position, whereas in the DNA photolyase, **T**<>**T** needs to move out of its original position in order to enter the active site. By studying a model with a dimer "flipped out" from the DNA helix, we decided to study a system which incorporated features of both the dinucleotide and the single-stranded nonamer systems described above. Initially, we constrained the dihedral angles on the deoxyribophosphate backbone around the **T**<>**T** lesion to cause a flip of the lesion out of its original position in the duplex (for definition of the dihedral angles and values chosen, see Supporting Information). This results in a single-stranded DNA system where the **T**<>**T** lesion is displaced from the position it would occupy in a normal duplex DNA. This system was then manually docked into the active site using the position of the **T**<>**T** dinucleotide as a guide. After restrained minimization, the restraints were removed, and the model was subjected to the MD protocol described above. The overall structure of the final enzyme−DNA complex is shown in Figure 5 in a slightly different orientation than the one shown in Figure 4. The 3′ side of the nonamer is again positioned to the right of the **T**<>**T** lesion, highlighted in gold.

Similar to those of the other two systems, interactions occur between the 5′ phosphate group and Arg^{342} and Arg^{397} and between the deoxyribophosphate group on the 3′ side of **T**<>**T** and Lys^{154}. Inside the active site, interactions occur between the carbonyl oxygen in the C4 position on the 3′ side and Trp^{384} and between the carbonyl oxygen in the C4′ position on the 5′ side of **T**<>**T** and Asp^{391}. In this system, **T**<>**T** lies intermediate to the other two systems. It is buried deeper in the active site than the **T**<>**T** in the base-stacked DNA system, as would be expected. However, the distance between the redox active FADH cofactor and **T**<>**T** is still larger than that in the dinucleotide−enzyme system. **T**<>**T** in the flipped out DNA−enzyme system is again shifted toward the 5′ direction as compared to the position of **T**<>**T** in the dinucleotide−enzyme system but not as far as the 5′ shift which occurred in the normal DNA−enzyme system.

The Enzyme−Double Strand DNA Model. Finally, we examined the interaction of a double-stranded DNA containing the **T**<>**T** lesion with DNA photolyase to test the assumption of a very similar binding of single-stranded and duplex DNA. For this purpose, the same 12 base pair duplex (GCGCTTAAGCGC) (CGCGAAT<>TCGCG) described above was chosen to allow comparisons with the results from earlier work[14,16,17] and from our validation studies. The **T**<>**T** lesion was flipped out by restraining the appropriate dihedral angles. Additional restraints were used to enforce base pairing on the 3′ side of the **T**<>**T** lesion. Several cycles of minimization and the gradual removal of restraints led to a duplex structure with reasonable base pairing and good base stacking on the strand opposite of the lesion. This structure was then superimposed on the single strand DNA−enzyme model and, after removal of the single strand, subjected to the minimization−MD−minimization protocol as described above. Figure 6 shows a side view of the final structure. In this representation, the 5′ side of the strand containing the **T**<>**T** dimer is shown on the right, and the 3′ side is on the left. The **T**<>**T** lesion, highlighted in yellow, is partially hidden in the active site.

The most noteworthy feature of this structure is the hole in the duplex, generated by flipping the dimer out of the position inside the duplex and into the active site. The base pairs on the opposite strand are unpaired, and the backbone at this position does not have any interactions with the enzyme. The overall

(33) For recent overviews, see: (a) Loyd, R. S.; Cheng, X. *Biopolymers* **1997**, *44*, 139−151. (b) Roberts, R. J.; Cheng, X. *Annu. Rev. Biochem.* **1998**, *67*, 181−198.

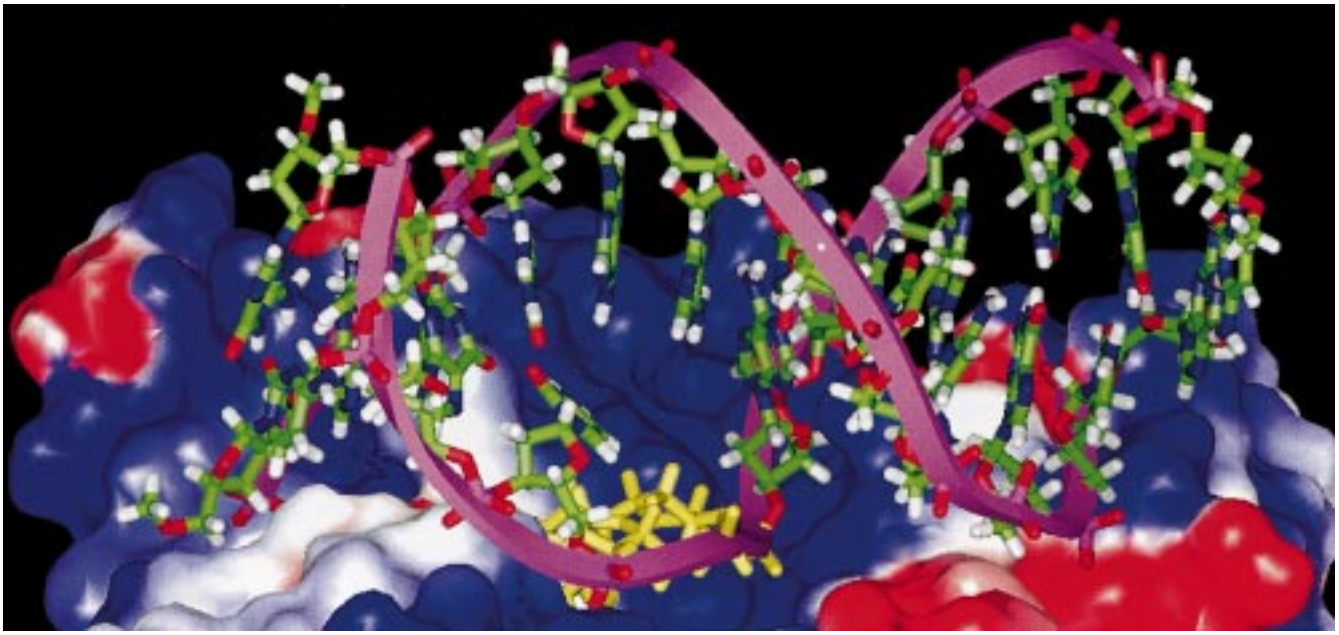

Figure 6. Complex of DNA photolyase with the approximate electrostatic potential projected onto the Connolly surface, complexed to a twelve base pair double-stranded DNA piece. The **T**<>**T** fragment, shown in yellow and partially hidden in the active site, is flipped out of the original position.

duplex structure is well-preserved, but the base pairing on the 3′ side of the lesion is altered. In contrast, base pairing on the 5′ side as well as the base stacking throughout the complementary strand is maintained. Similar to the single-strand models discussed above, the strand containing the **T**<>**T** lesion does not have strong interactions with the enzyme on the 5′ side beyond the immediate region of the lesion. These results are in agreement with the experiments of Hussain et al., which gave very similar binding constants for single- and double-stranded DNA,[20] thus indicating that there are very few or no contacts between the enzyme and the complementary strand. It is not clear at this point whether the alteration of the base pairing on the 3′ side of the lesion is entirely due to the strong interactions between the duplex and the solvent-exposed amino acids or if they are in part due to the relatively short simulation time. However, the finding that the interactions on the 3′ side lead to an alteration of the base pairing, whereas no significant changes in the base stacking or pairing are obtained on the 5′ side of the lesion, is in line with the experimentally observed differences of the effect of alkylations on the binding constants. Although the experimental and computational results are in good agreement, the solution structure and dynamics of the uncomplexed duplex with the flipped-out **T**<>**T** lesion need to be studied in more detail to obtain further insight into the effect of base flipping on the structure of DNA and the efficiency of DNA repair.[34]

The relative position of the **T**<>**T** in the different model systems is summarized in Figure 7, which shows a detailed view of the active site with an overlay of the positions of the dimer in the dinucleotide−photolyase complex (shown in red), the position of the dimer in the flipped-out single strand DNA-photolyase complex (shown in cyan), the position of the dimer in the flipped-out duplex DNA-photolyase complex (shown in yellow), and the position of the dimer in the single strand DNA-photolyase complex (shown in blue). In comparing the dimer−enzyme, DNA−enzyme, and the two flipped out DNA−enzyme

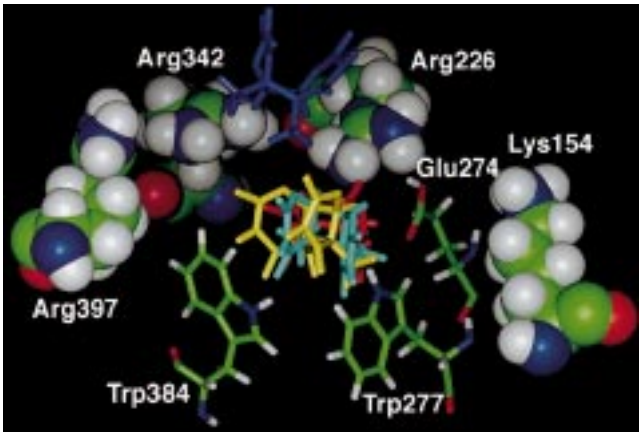

Figure 7. Overlay of the relative positions of the **T**<>**T** models of the enzyme−substrate complexes of DNA photolyase with the dinucleotide−photolyase complex (red), the flipped-out single strand DNA-photolyase complex (cyan), the flipped-out duplex DNA-photolyase complex (yellow), and the single strand DNA-photolyase complex (blue). The deoxyribophosphate backbone has been removed for clarity.

systems, the dimers in the flipped out systems lie between the positions of the other two systems both in its height and in its 5′ shift. It is noteworthy that the position of the two dimers in the flipped-out single-stranded DNA and in the duplex DNA occupy very similar positions, reemphasizing the experimentally observed similarities in the mode of binding in the two systems. The **T**<>**T** dimers in the flipped-out systems are also turned clockwise between 15 and 20° from the other two dimers, a position which appears to serve as an intermediary position. They are closer to the tryptophans than the dimers from either the dinucleotide−photolyase complex or the DNA-photolyase complex with the **T**<>**T** in its base-stacked position. Concurrently, the entire strand appears to lean away from the tryptophan side of the active site, so that the strand lies further away from the active site than the normal DNA strand, and this forces the dimer closer to the tryptophans.

(34) For recent experimental work in the area, compare: Butenandt, J.; Burgdorf, L. T.; Carell, T. *Angew. Chem., Int. Ed. Engl* **1999**, *38*, 708−711.

Comparison with DNA Photolyase from *Anacystis nidulans*. After the completion of this study, the three-dimensional structure of a DNA photolyase from the cyanobacterium *Anacystis nidulans* became available.[35] Even though this photolyase is a member of the deazaflavin class, the overall structure is nevertheless very close to the *E. coli* photolyase. Like the active site in most photolyases, the active site of this enzyme shows a particularly high degree of structural similarity, with only one out of sixteen amino acids interacting with the FADH cofactor changed. We anticipate that the conclusions drawn from our model of the *E. coli* enzyme−substrate complex are equally applicable here. The almost superimposable relative positions of the FADH, Tyr^{290}, Trp^{286}, and Trp^{390} (equivalent to Tyr^{281}, Trp^{277}, and Trp^{384} in *E. coli* photolyase) again underlines the importance of these amino acid residues for the function of the enzyme. This would be expected if the π-systems of the aromatic amino acid side chains would be essential for electron transfer. It therefore appears that the proposed model is in good agreement with the three-dimensional structure of the DNA photolyase of *Anacystis nidulans*.

Summary and Conclusion

In this paper, we present a model for the enzyme−substrate complex of *E. coli* DNA photolyase with several dinucleotides and single-stranded and duplex DNA fragments containing cyclobutane pyrimidine dimers **T<>T**. These models provide structural information to explain previous experimental findings. The different effects of ethylation of the backbone of the DNA in the 3′ and 5′ direction correspond to the pronounced directionality of binding with the DNA−protein interactions in our model, extending much further on the 3′ side. The different rates of electron transfer for various cyclobutane pyrimidine dinucleotides are rationalized by the steric interaction of the methyl groups in the 5 position of the dimer with Arg^{226} and Arg^{342}.

The proposed model rationalizes the substrate recognition through a combination of nonspecific ion-pair interactions between the solvent-exposed, positively charged residues on the surface of the enzyme and the deoxyribophosphate backbone of the DNA. Within the active site, specific polar contacts and hydrogen bonds occur. One side of the active site is rather nonpolar with aromatic amino acid side chains making hydrophobic contacts with the cyclobutane moiety of **T<>T**. The results from the calculations and the high degree of conservation of these aromatic amino acids also suggest that they have an important function in preventing an electrophilic attack on the radical anion intermediate of the reaction and in mediating the electron transfer from the redox active FADH cofactor through their π-systems. Site-directed mutagenesis studies of Trp^{277}, Tyr^{281}, and Trp^{384} and determination of the binding constants and quantum yields for repair by these mutants could provide experimental evidence for these hypotheses.

Acknowledgment. We gratefully acknowledge financial support of this research by the National Institutes of Health (Grant CA73775-01A1) and the generous allocation of computing resources by the Office of Information Technology at the University of Notre Dame and the National Center of Supercomputer Applications (NCSA). D.B.S. thanks the Howard Hughes Summer Research Program and the Barry M. Goldwater Foundation for undergraduate fellowships.

Supporting Information Available: AMBER parameters used for the cofactors and the thymine dimer and the coordinates of all structures discussed (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA982660Y

(35) (a) Miki, K.; Tamada, T.; Nishida, H.; Inaka, K.; Yasui, A.; deRuiter, P. E.; Eker, A. P. M. *J. Mol. Biol.* **1993**, *233*, 167−169. (b) Tamada, T.; Kitadokoro, K.; Higuchi, Y.; Inika, K.; Yasui, A.; de Ruiter, P. E.; Eker, A. P. M.; Miki, K. *Nat. Struct. Biol.* **1997**, *4*, 887−891.