Theoretical Study of Electron Transfer between the Photolyase Catalytic Cofactor $FADH^-$ and DNA Thymine Dimer

Jens Antony, Dmitry M. Medvedev, and Alexei A. Stuchebrukhov*

Contribution from the Department of Chemistry, University of California, Davis, California 95616

Received October 22, 1999

Abstract: Photolyase is an enzyme that catalyzes photorepair of thymine dimers in UV damaged DNA by electron-transfer reaction. We docked a thymine dimer to photolyase catalytic site, using crystal structure coordinates of the substrate-free enzyme from *Escherichia coli*, studied molecular dynamics of the system, and calculated the electron-transfer matrix element between the lowest unoccupied molecular orbitals of flavin and the dimer. We find that the rms transfer matrix element along the dynamic trajectory is about 6 cm^{-1}, which is consistent with the experimentally determined rate of transfer. In the average configuration the docked thymine dimer is sitting deep in the catalytic site, and approaches the adenine of FAD with the C4═O4 carbonyl groups. The average distance between the flavin and the base pair is less than 3 Å. The electron-transfer mechanism utilizes the unusual conformation of FAD in photolyases, in which the isoalloxazine ring of the flavin and the adenine are in close proximity, and the peculiar features of the docked orientation of the dimer. The calculations show that despite the short distance between the donor and acceptor complexes, the electron-transfer mechanism between the flavin and the thymine bases is not direct, but indirect, with the adenine acting as an intermediate.

1. Introduction

Far-UV light is harmful to genetic stability of DNA in which it can produce several different types of photochemical modifications.[1] The dimerization of two adjacent pyrimidine bases is the most common defect. DNA photolyases are enzymes that catalyze photorepair of DNA. These proteins bind to the damaged part of DNA and cause its repair when exposed to UV−visible light (photoreactivation). A central step of this repair mechanism is the transfer of an electron from one of the two cofactors of DNA photolyase, an excited flavin molecule ($*FADH^-$), to the pyrimidine dimer. As a consequence of this electron-transfer reaction, the two bonds between the pyrimidine bases break up. Upon dimer splitting the electron is transferred back to the oxidized flavin, and the enzyme and the repaired DNA dissociate. The present knowledge of the mechanism of light-induced repair of DNA by photolyase is summarized in several review articles.[2−4]

The crystal structures of photolyase from *Escherichia coli*[5] and *Anacystis nidulans*[6] have been recently resolved. However the structure of DNA/photolyase complex is unknown at present. Several amino acid residues that are involved in different aspects of the enzymatic function have been tested by site-directed mutagenesis.[7−9] Substrate−enzyme binding has been studied for native[10] and synthetic[11,12] substrates. It was found that the enzyme binds specifically to UV-irradiated DNA in a sequence-independent way regardless of whether the DNA is in the superhelical, open circular, or linear form or whether the DNA is single or double stranded. This latter result strongly suggests that the enzyme predominantly recognizes the dimer rather than any distortion of the helix.[10]

The repair reaction involves the catalytic cofactor in photolyase, $FADH^-$. The first excited singlet state $*FADH^-$, formed by a direct absorption of a photon or by energy transfer from the second chromophore (MTHF in the folate and 8-HDF in the deazaflavin class), initiates the splitting of Pyr<>Pyr by electron transfer. The natural substrate for photolyase is Pyr<>Pyr in a duplex DNA. Even though with smaller affinity, however, photolyase also binds to dinucleotide thymine dimers and splits them with the same quantum yield as in the oligonucleotide form.[13]

The rate and efficiency of single electron transfer from $*FADH^-$ to Pyr<>Pyr have been investigated by time-resolved fluorescence, absorbance,[14−18] and EPR spectroscopies.[19] The

* Address correspondence to this author.

(1) Friedberg, E. C.; Walker, G. C.; Siede, W. *DNA Repair and Mutagenesis*; ASM Press: Washington, DC, 1995.
(2) Heelis, P. F.; Hartman, R. F.; Rose, S. D. *Chem. Soc. Rev.* **1995**, *24*, 289.
(3) Sancar, A. *Biochemistry* **1994**, *33*, 2.
(4) Kim, S.-T.; Sancar, A. *Photochem. Photobiol.* **1993**, *57*, 895.
(5) Park, H.-W.; Kim, S.-T.; Sancar, A.; Deisenhofer, J. *Science* **1995**, *268*, 1866.
(6) Tamada, T.; Kitadokoro, K.; Higuchi, Y.; Inaka, K.; Yasui, A.; de Ruiter, P. E.; Eker, A. P. M.; Miki, K. *Nat. Struct. Biol.* **1997**, *4*, 887.
(7) Kim, S.-T.; Li, Y. F.; Sancar, A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 900.
(8) Li, Y. F.; Heelis, P. F.; Sancar, A. *Biochemistry* **1991**, *30*, 6322.
(9) Li, Y. F.; Sancar, A. *Biochemistry* **1990**, *29*, 5698.
(10) Sancar, G. B.; Smith, F. W.; Sancar, A. *Biochemistry* **1985**, *24*, 1849.
(11) Husain, I.; Sancar, A. *Nucleic Acids Res.* **1987**, *15*, 1109.
(12) Husain, I.; Sancar, G. B.; Holbrook, S. R.; Sancar; A. *J. Biol. Chem.* **1987**, *262*, 13188.
(13) Kim, S.-T.; Sancar, A. *Biochemistry* **1991**, *30*, 8623.
(14) Langenbacher, T.; Zhao, X.; Bieser, G.; Heelis, P. F.; Sancar, A.; Michel-Beyerle, M. E. *J. Am. Chem. Soc.* **1997**, *119*, 10532.
(15) Kim, S.-T.; Volk, M.; Rousseau, G.; Heelis, P. F.; Sancar, A.; Michel-Beyerle, M.-E. *J. Am. Chem. Soc.* **1994**, *116*, 3115.
(16) Kim, S.-T.; Heelis, P. F.; Sancar, A. *Biochemistry* **1992**, *31*, 11244.
(17) Kim, S.-T.; Heelis, P. F.; Okamura, T.; Hirata, Y.; Mataga, N.; Sancar, A. *Biochemistry* **1991**, *30*, 11262.
(18) Okamura, T.; Sancar, A.; Heelis, P. F.; Begley, T. P.; Hirata, Y.; Mataga, N. *J. Am. Chem. Soc.* **1991**, *113*, 3143.
(19) Kim, S.-T.; Sancar, A.; Essenbacher, C.; Babcock, G. T. *J. Am. Chem. Soc.* **1992**, *114*, 4442.

10.1021/ja993784t CCC: $19.00
Published on Web 01/28/2000

rate of the primary electron transfer is found to depend on the nature of the pyrimidine dimer.[14] In *E. coli* photolyase the rate at 275 K is 10^{10} s^{-1} for the case of a thymine−thymine dimer and three times faster for a uracil−uracil dimer. The mixed dimer T<>U acts like the symmetric thymine dimer, and the U<>T like the symmetric uracil dimer.[14]

The similarity of quantum yields for the dinucleotide and oligonucleotide forms of a pyrimidine dimer implies that the dimer has the same position in the enzymes active site in both cases.[13] Under this assumption it is possible to separate DNA binding and dimer recognition by photolyase from the electron-transfer process by considering a pyrimidine dimer dinucleotide as the substrate. With knowledge of the asymmetry in structure and polarity of the binding pocket from the crystal structure[5,6] and the measured rates of electron transfer,[14,16,17] conclusions concerning the binding geometry of the pyrimidine dimer should be possible.

In recent works the binding of DNA photolyase to model pyrimidine dimer dinucleotides[20] and oligonucleotides[21] was studied with molecular dynamics simulations. The simulated enzyme−substrate complexes showed no close contacts between the damaged base pair and the FAD cofactor. The minimum van der Waals contact is 5 Å for a bare U<>T dimer and 9 Å for the dinucleotide.[20] The closest contact reported in ref 21 is 6 Å and occurs at the ribose side chain of the catalytic cofactor. Consequently an electron-transfer pathway mediated by the π-systems of Tyr281 and Trp277 or by Trp384 was proposed.

In this paper, we address the question of the docking configuration of the thymine dimer and the mechanism of electron transfer with a different theoretical approach. We calculate the electronic coupling matrix element between the flavin molecule and the dimer bound in the active site of the enzyme and evaluate the likelihood of a particular model of binding by comparing the maximum electron-transfer rate possible for the coupling obtained with experimental data.[14] The DOCK 4.0 program[22,23] is used for the prediction of global features of the bound configuration of the dimer. Then, molecular dynamics simulation is used to study dynamic behavior of the enzyme−dimer complex obtained in the docking procedure. Water is included in the simulation. The coupling matrix element is calculated and averaged along the dynamic trajectory. The simulation is done for all four combinations of thymine and uracil bases in order to compare our results with the results of a recent experiment[14] that revealed a weak dependence of the reaction rate on the nature of dimer.

We find that the electron-transfer mechanism utilizes the unusual conformation of FAD in photolyases, which brings the isoalloxazine ring of the flavin and adenine in close proximity,[5,6,24] and the peculiar features of the docked orientation of the dimer. The calculations show that despite the close proximity between the donor and acceptor complexes, the electron-transfer mechanism between the flavin and the thymine bases is not direct, but indirect, with the adenine acting as an intermediate. Water plays significant role in the process by modifying the dynamic behavior of the system.

The rms transfer matrix element along the molecular dynamics trajectory is about 6 cm^{-1}, which is consistent with the experimentally determined rate of transfer, while configurations with maximal coupling have the values of the matrix element about 30 cm^{-1}. In the average configuration the docked thymine dimer is sitting deep in the catalytic site, with the distance between the flavin and the base pair less than 3 Å. This configuration resembles the one that was recently proposed by Park et al.[5]

On the basis of these results and our study of rigid body docking modeling of the DNA/photolyase complex (data not presented), we conclude that, in order for a T<>T dimer built in a DNA duplex to assume the position found for the model T<>T dinucleotide, a significant deformation of the damaged strand in the DNA/photolyase complex is necessary. This finding supports the recently supplied evidence to the proposal that the T<>T dimer flips out of the double-stranded DNA structure upon photolyase binding before the repair redox chemistry occurs[25] and the results of a recent molecular dynamics simulation study.[21]

The paper is structured as follows. In the next section, the methods of the determination of the docked orientation, molecular dynamics simulations, and calculation of the transfer matrix element are described. In section 3, docking configurations are discussed, donor and acceptor states are identified, and estimated values for the coupling constants are given as function of the docked configuration and the donor and acceptor orbitals. Results of molecular dynamics simulations and calculation of matrix element along the molecular dynamics trajectories are also presented. Finally, implications of the results obtained and open questions are indicated.

2. Methods

2.1. Docking. The X-ray structure of the pyrimidine dimer−photolyase complex is not known. We model the binding geometry of a dimer in the active site of DNA photolyase using a combination of the docking procedure with DOCK 4.0 program[22,23] and molecular dynamics simulation.

In the docking procedure, the crystallographic coordinates of the chain A of the enzyme[5] without water and those of a thymine dimer obtained with a molecular dynamics simulation[26] of a piece of damaged DNA were used.

The DOCK program generates many possible orientations of the dimer within the catalytic site of the enzyme. These configurations are scored using several schemes designed to measure steric and/or chemical complementarity of the receptor−ligand complex. The scores are used to evaluate the likelihood of a given orientation of the dimer in the catalytic site.

Specifically, the DOCK program matches the shape of the ligand with that of the receptor's binding site. The active site is composed of the amino acids of DNA photolyase in the vicinity of the catalytic cofactor FAD, i.e., of all amino acids that have at least one heavy atom within a given radius of any heavy atom of FAD. The shape of the active site is determined by the water accessible surface of this part of the protein.[27,28] DOCK models the probable binding sites, i.e., the concave parts of this surface, by filling them with mutually overlapping spheres. The centers of these spheres constitute possible positions of ligand atoms in the binding site. Docking consists of finding reasonable pairings between ligand atoms and sphere centers and orienting the ligand in the binding site in a way that the root-mean-square (rms) distance of all pairs is minimized. The docking procedure assumes that the receptor is rigid. For details of the procedure we refer to the papers written on DOCK, e.g. the refs 29 and 30 and the DOCK manual.[31]

(20) Hahn, J.; Michel-Beyerle, M.-E.; Rösch, N. *J. Phys. Chem. B* **1999**, *103*, 2001.
(21) Sanders, D. B.; Wiest, O. *J. Am. Chem. Soc.* **1999**, *121*, 5127.
(22) Kuntz, I. D.; Meng, E. C.; Shoichet, B. K. *Acc. Chem. Res.* **1994**, *27*, 117.
(23) Kuntz, I. D. *Science* **1992**, *257*, 1078.
(24) Hahn, J.; Michel-Beyerle, M.-E.; Rösch, N. *J. Mol. Model.* **1998**, *4*, 73.
(25) Berg, B. J. V.; Sancar, G. B. *J. Biol. Chem.* **1998**, *273*, 20276.
(26) Miaskiewicz, K.; Miller, J.; Cooney, M.; Osman, R. *J. Am. Chem. Soc.* **1996**, *118*, 9156.
(27) Connolly, M.-L. *Science* **1983**, *221*, 709.
(28) *Program No. 429 MS*; Quantum Chemical Program Exchange: La Jolla, CA, Dec 1983.
(29) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. *J. Mol. Biol.* **1982**, *161*, 269.

In addition to rms distance score, DOCK calculates the nonbonded energy as a sum of the Coulombic and van der Waals interactions, using a precomputed grid in the vicinity of the binding site.[32,33] For this purpose, hydrogen atoms have been added to the coordinate files and partial charges have been assigned using the InsightII program[34] assuming pH = 7. The protonation results in negatively charged aspartate and glutamate and positively charged lysine and arginine residues. All other amino acids are neutral. For docking procedure ESFF[35] force field was used.

The primary purpose of the docking procedure is to obtain a correct global orientation of the dimer in the catalytic site. This configuration is used then as a starting point for the following molecular dynamics simulation. It is important to notice that we are looking for configurations that lead to the observed rate of electron transfer between the flavin cofactor and the dimer. Therefore the scoring of various configurations includes not only the binding energy but also the strength of electronic coupling. The main idea of our docking analysis is to look among the configurations that have high binding energy for the one that leads to the largest transfer matrix element. In this way we find the global orientation of the dimer. The docking model then is further improved in the molecular simulation, as described below. The unusually high electron-transfer rate[14,16,17] of the order of 10^{10} s^{-1}, the efficiency[13] of electron transfer, and the fact that the dimer in the binding pocket has no other role than to accept an electron make us believe that the enzyme binds the substrate in a way that optimizes the electron-transfer rate.

2.2. Molecular Dynamics Simulation. Following the docking procedure that determined the global orientation of the dimer in the catalytic site, the molecular dynamics (MD) simulation of the system was performed with the simulation package Amber 5.0.[36,37] The protein was substituted by its active site (see subsection 3.6). In all stages of the simulation, the movable parts were the dimer, FAD in the binding pocket, and the added water molecules. Other enzyme atoms were not allowed to move. The initial structure was the one obtained in the docking procedure.

The molecular dynamics simulation was performed using the Amber94 force field.[38] The MD force field requires partial charges to be assigned to the atoms of the system. The partial charges to the atoms of protein and of water molecules were assigned automatically by the xLEaP program (part of Amber 5.0). For dimer and FAD an ab initio quantum chemical calculation was performed, using the Gaussian94 package.[39] These calculations were performed with the restricted Hartree−Fock method using the 6-31G* basis set. The charges for MD simulations were calculated for singlet ground states of both $FADH^-$ and thymine dimer. Since Mulliken charges do not reproduce electrostatic potentials of molecules very well,[36] additional charge fitting procedure was performed using the program RESP,[40,41] which is a part of Amber 5.0, as recommended.[36]

Water was added as follows. At first, the dimer was taken out of the structure and water molecules were added inside and around the binding pocket. Minimization of potential energy of this system was performed, during which the bonds not involving hydrogen atoms were relaxed. After the minimization step, the equilibration of the system was performed, starting with zero kinetic energy and ending with the kinetic energy of atoms corresponding to temperature $T = 300$ K. Then a molecular dynamics run with constant temperature $T = 300$ K was performed. During the equilibration and molecular dynamics runs all bonds in the system were constrained.

After the equilibration and MD simulation of water in the catalytic site, the dimer was added to the system. It was positioned as predicted by the docking procedure, and those water molecules that had significant overlap with the atoms of the dimer were removed.

After addition of the dimer to the catalytic site containing water, as described above, the energy minimization was performed again. Here again, the bonds that do not involve hydrogen atoms were relaxed. The structure obtained as a result of the energy minimization was used then as the starting point for the equilibration run. From this stage on, all bonds in the system were constrained.

At first, the equilibration lasting 100 ps was performed, starting with zero kinetic energy and finishing with kinetic energy of atoms corresponding to temperature $T = 300$ K. Following the last equilibration procedure, the main molecular dynamics run with constant temperature $T = 300$ K was performed. It lasted 1 ns.

The procedure outlined above was carried out for all four dimers: T<>T; T<>U; U<>T; U<>U. The partial charges for all dimers were taken to be equal to those for T<>T, and the partial charge on the hydrogen atoms that replaced the methyl groups of the thymines were taken to be equal to the charges on these methyl groups.

2.3. Electronic Coupling. In this calculation, the electronic structure of the protein was described within the semiempirical extended Hückel method.[42] The states of the valence electrons are represented in this method by single-ζ Slater-type atomic orbitals

$$\phi = Nr^{n-1}e^{-\zeta r}Z_{lm} \tag{1}$$

Here n is the main quantum number, ζ is the Slater exponent, Z_{lm} is the angular part for a given angular momentum l and magnetic quantum number m, N is the normalization constant, and r is the distance from the atomic nucleus. The diagonal elements of the Hamiltonian are the empirical valence shell ionization potentials. The nondiagonal matrix elements are given by the Wolfsberg−Helmholtz expression

$$H_{ij} = KS_{ij}\frac{H_{ii} + H_{jj}}{2} \tag{2}$$

with $K = 1.75$. Slater exponents and valence shell ionization potentials are taken from the ref 43 and are listed in Table 1.

The coupling between donor and acceptor electronic states $|D\rangle$ and $|A\rangle$ with overlap S_{DA} is given by the matrix element[44,45]

$$T_{DA} = \frac{H_{DA} - ES_{DA}}{1 - {S_{DA}}^2} \tag{3}$$

(30) Shoichet, B. K.; Bodian, D. L.; Kuntz, I. D. *J. Comput. Chem.* **1992**, *13*, 380.
(31) DOCK WWW page. http://www.cmpharm.ucsf.edu/kuntz/dock.html (accessed Sept 22, 1997).
(32) Meng, E. C.; Gschwend, D. A.; Blaney, J. M.; Kuntz, I. D. *Proteins* **1993**, *17*, 266.
(33) Meng, E. C.; Shoichet, B. K.; Kuntz, I. D. *J. Comput. Chem.* **1992**, *13*, 505.
(34) *InsightII 4.0.0 Molecular Modeling System User Guide*; Molecular Simulations Inc.: San Diego, CA, Sept 1996.
(35) *Discover 2.9.8/96.0/4.0.0 Force field Simulations User Guide*; Molecular Simulations Inc.: San Diego, CA, Sept 1996.
(36) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Ferguson, D. M.; Radmer, R. J.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 5*; University of California: San Francisco, CA, 1997.
(37) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E., III; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1.
(38) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.;, Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(39) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94, Revision E.2*; Gaussian, Inc.: Pittsburgh, PA, 1995.
(40) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269.
(41) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620.
(42) Yates, K. *Hückel Molecular Orbital Theory*; Academic Press: New York, 1978.
(43) Alvarez, S. *Tables of Parameters for Extended Hückel Calculations*; Universitat de Barcelona: Barcelona, Spain, Dec 1995.
(44) Daizadeh, I.; Gehlen, J. N.; Stuchebrukhov, A. A. *J. Chem. Phys.* **1997**, *106*, 5658.
(45) Newton, M. D. *Chem. Rev.* **1991**, *91*, 767.

Table 1. Parameters for Extended Hückel Calculation[43] with Valence Shell Ionization Potentials (H_{ii}) in eV and Slater Exponents (ζ) in au

		H_{ii}		ζ	
atom	shell n	s	p	s	p
H	1	−13.6		1.3	
C	2	−21.4	−11.4	1.625	1.625
N	2	−26.0	−13.4	1.950	1.950
O	2	−32.3	−14.8	2.275	2.275
P	3	−18.6	−14.0	1.75	1.30
S	3	−20.0	−11.0	2.122	1.827

T_{DA} is the half-energy splitting between two delocalized eigenstates $|\psi_+\rangle$ and $|\psi_-\rangle$ which result from diagonalization of H in a configuration when the states $|D\rangle$ and $|A\rangle$ are in resonance and have common energy E. When there is no resonance, $|D\rangle$ and $|A\rangle$ states can be approximated by two corresponding eigenstates of the entire system. These two eigenstates are identified by projecting all eigenstates of the system onto zeroth-order donor and acceptor states $|d\rangle$ and $|a\rangle$, which are obtained by diagonalization of the isolated donor and acceptor complexes, respectively. The donor and acceptor complexes are predefined sets of atoms, on which the states $|D\rangle$ and $|A\rangle$ are likely to be localized. Usually, the overlap of only one of the eigenstates of the entire system with zeroth-order state $|d\rangle$ or $|a\rangle$ is close to unity, while all other eigenstates have nearly zero overlaps. The interaction of the zeroth-order donor and acceptor states with the protein environment results only in their slight delocalization. However, this delocalization causes the effective interaction between the donor and acceptor states.

Since the effective coupling between the donor and acceptor states is of the order of 1−30 cm^{-1} and small on the scale of molecular orbital energies (order of 1 eV), the nuclear configurations of the system derived from its crystal structure or from MD simulation are unlikely to be in a transition state fulfilling the resonance condition of electron tunneling. In electron-transfer reactions, the donor and acceptor states are brought into resonance by the motion of the protein atoms and fluctuations of the solvent. Here we model the former effect by adding energies to the Hamiltonians of the donor or/and acceptor complexes. The additional energies give rise to a shift of the electronic spectra of the donor and acceptor complexes. In particular, states $|D\rangle$ and $|A\rangle$ are shifted and eventually brought into resonance. The splitting between delocalized states $|\psi_+\rangle$ and $|\psi_-\rangle$ formed from $|D\rangle$ and $|A\rangle$ at the transition state gives the accurate value of electronic coupling.[46]

The search for the transition state in ET reaction can be computationally costly, since it involves multiple diagonalization of the whole protein (or at least a large representative part of it). An alternative method to calculate the electronic coupling, which requires only one diagonalization, is the method of tunneling currents.[47−49] No precise resonance between $|D\rangle$ and $|A\rangle$ is needed in this method. Instead it is sufficient to adjust their zeroth-order approximations, $|d\rangle$ and $|a\rangle$. In this method the matrix of interatomic tunneling currents is defined by

$$J_{pq} = \sum_{i\in p}\sum_{j\in q} J_{pi,qj} = \frac{1}{\hbar}\sum_{i\in p}\sum_{j\in q}(H_{pi,qj} - ES_{pi,qj})(c_{qj}^{D}c_{pi}^{A} - c_{qj}^{A}c_{pi}^{D}) \quad (4)$$

Here c^D and c^A are the molecular orbital expansion coefficients of the donor and acceptor states, respectively, and E is the tunneling energy, which was taken to be equal to average of the energies of donor and acceptor states. The elements $J_{pi,qj}$ and $J_{p,q}$ describe the probability flux between atomic orbitals $|pi\rangle$ and $|qj\rangle$ and the respective atoms p and q during electron tunneling process. The tunneling matrix element is related to the total tunneling flux through a dividing surface S_D by

$$T_{DA} = -\hbar \max_{S_D}\left(\sum_{p\in S_D}\sum_{q\notin S_D} J_{pq}\right) \quad (5)$$

(46) Katz, D. J.; Stuchebrukhov, A. A. *J. Chem. Phys.* **1998**, *109*, 4960.
(47) Stuchebrukhov, A. A. *J. Chem. Phys.* **1997**, *107*, 6495.
(48) Stuchebrukhov, A. A. *J. Chem. Phys.* **1996**, *105*, 10819.
(49) Stuchebrukhov, A. A. *J. Phys. Chem.* **1996**, *100*, 8424.

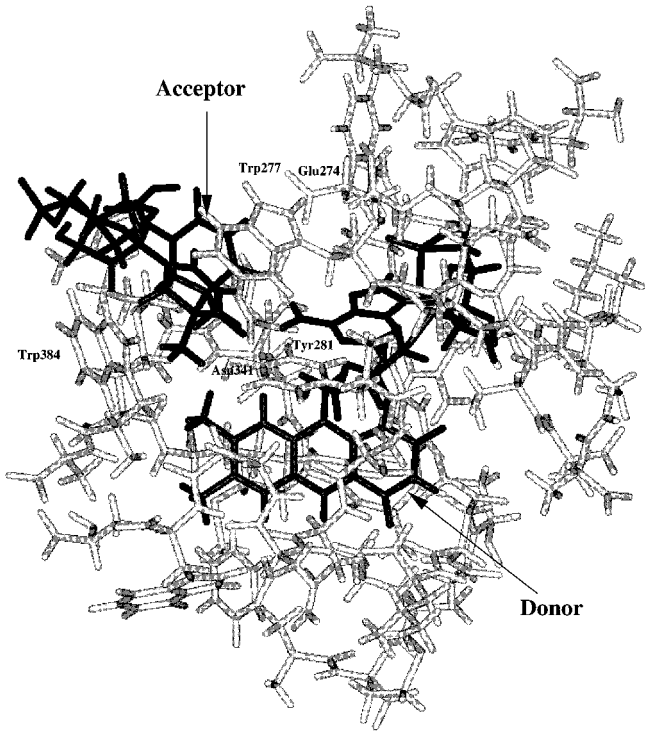

Figure 1. The structure found in docking.

The surface here is chosen as a sphere around the donor complex, with radius sufficiently large so that it includes most of the electron density in donor state but excludes most of electron density in acceptor state.

3. Results and Discussion

3.1. Docking. The global orientation of the dimer in the catalytic site of photolyase obtained with the docking procedure is shown in Figure 1. In this section we describe the details of this configuration and the way how it was obtained.

First, to assess the performance of DOCK on our system, we removed the catalytic cofactor FAD from the crystal structure and redocked it to the apoenzyme. We find that the DOCK program can indeed find a correct position of the FAD in the enzyme pocket, provided that the spheres that model the binding site and other input parameters are chosen appropriately. The energy score, however, was found to be an insufficient measure for the quality of a docking orientation, because the receptor in DOCK is assumed to be rigid and the interaction energy is estimated only in a rough manner. We concluded that to determine the correct orientation when the docking is done on an unknown structure, additional guiding information is needed. In our case the electronic coupling strength was used for this purpose.

The experiments[14,16,17] show that electron transfer to a dimer bound in the pocket occurs with a very high rate. This indicates that the distance r_{df} between the dimer and the flavin molecule of FAD, where the donor wave function is localized, should be equal, or close, to that of van der Waals contact.

Among the likely configurations that we determined with the rigid ligand receptor docking, the distance between the dimer and the flavin, r_{df}, ranges from 2.5 to 5.5 Å. Most of orientations, in which the dimer approaches the flavin within less than 3 Å, have a positive energy score. We find, however, that for some orientations the repulsion originates only from a few unfavorable van der Waals overlaps. For a flexible ligand in a flexible receptor, only a slight deformation would be necessary to accommodate the dimer into almost the same orientation without any atomic overlaps.

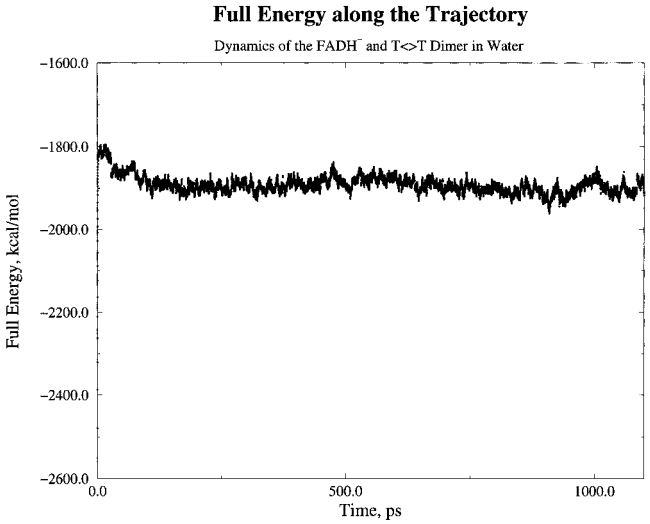

Figure 2. Full energy along the dynamics trajectory.

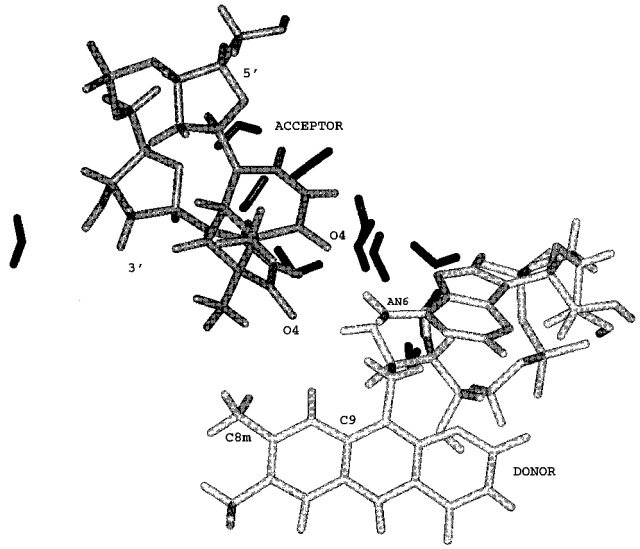

Figure 3. Characteristic geometry of redox cofactors in the MD simulation.

For all orientations with r_{df} below 3 Å the atoms involved in the closest contact are two hydrogens of the C8m methyl group of the flavin and the methyl group of the 3′ thymine (for notation see Figure 3). We find that once the thymine dimer approaches the flavin to almost van der Waals contact, its orientation in the binding site becomes more or less fixed. For larger intermolecular distances, r_{df}, the dimer can assume many orientations. Interestingly, in all these configurations the closest contact with FAD is established by the 3′ thymine methyl group, while the electron-transfer rate is affected by the 5′ thymine methyl group of the dimer.

Using both rigid and flexible docking and utilizing different options provided by DOCK, we finally found a structure which approaches the flavin within 2.8 Å and has the largest coupling strength. This structure, shown in Figure 1, was used as the starting point for the MD simulation.

3.2. Molecular Dynamics Simulation. Molecular dynamics simulation was performed as described in the Methods section.

The behavior of the total energy of the system with T<>T dimer in the binding site along the dynamics trajectory is shown in Figure 2. The first 100 ps represents the equilibration procedure, and next 1000 ps the main molecular dynamics run. The graphs of total energy along the dynamics trajectories of systems with three other dimers are similar.

(50) Kim, S.-T.; Sancar, A.; Essenbacher, C.; Babcock, G. T. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8023.

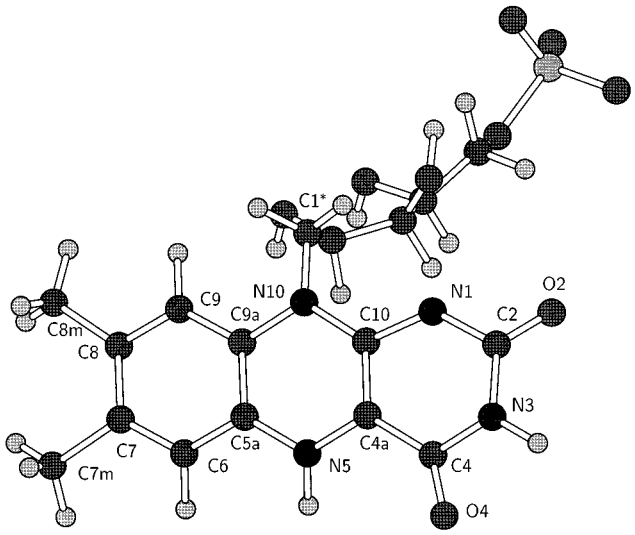

Figure 4. Geometric structure of the donor complex.

A characteristic relative orientation of redox cofactors and the nearest water molecules in the molecular dynamics simulation is shown in Figure 3. The average position of the dimer along the dynamic trajectory slightly shifts out of the FAD pocket, compared with the structure found in docking, Figure 1. The average distances between the oxygens of C4═O4 carbonyl groups and the AN6 atom of adenine are now about 2.9 Å with variance of about 0.15 Å (5′ thymine) and 3.9 Å with variance of about 0.30 Å (3′ thymine), while in the initial structure they are equal to 2.4 and 2.6 Å, respectively. For other three dimers the distances between the oxygen of the 5′ C4═O4 carbonyl group and the AN6 atom of adenine are approximately the same, while the distances between the oxygen on the 3′ C4═O4 carbonyl group and the AN6 atom of adenine differ by values of up to 1.4 Å. The average closest distance between T<>T dimer and FAD is now about 2.5 Å, and the closest contact is established between the hydrogen atoms on the C8m methyl group of the FAD and those on the methyl group of the 3′ thymine.

Most of the added water molecules are located outside of the binding pocket, and only few of them get inside the pocket. No water molecules were found right between the dimer and FAD, but some of them are located rather close to the dimer—in most of the structures obtained in MD simulations, from 5 to 8 water molecules were found within 7 Å of the oxygen of C4═O4 carbonyl group on 3′ thymine. For other three dimers similar results were obtained.

3.3. Electronic Structure of FAD. The catalytic cofactor of photolyase is a flavin adenine dinucleotide[3] (FAD). It is bound to the enzyme in a U-shaped conformation, opening a hole on the flat surface of the helical domain of the protein.[5] To be catalytically active, the cofactor has to be in the reduced form $FADH^-$ and in its excited singlet state.[50] Here we describe the electronic structure of the system in extended Hückel approximation. For the geometry we use the crystal structure data of DNA photolyase from *E. coli*,[5] with hydrogen atoms added by the InsightII program[34] (see Figure 4).

The highest occupied and lowest unoccupied molecular orbitals of FAD are localized either on the flavin or on the adenine and resemble closely the corresponding states calculated for each entity separately (see Figure 5). The highest occupied

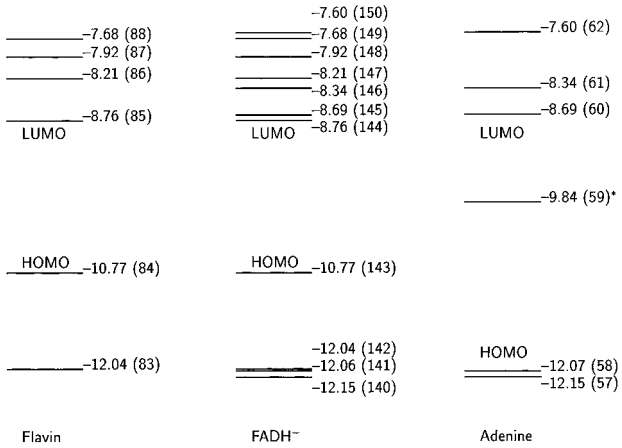

Figure 5. Electronic structure of FAD in the extended Hückel approximation for the crystal structure geometry with correlation diagram of the highest occupied and lowest unoccupied molecular orbitals for the whole cofactor with those for the part containing the flavin and the adenine. Orbital energies are given in eV; orbital numbers (in parentheses) start with zero. The 59 state on the adenine part has 99% of its amplitude on the phosphate group and is therefore an artificial state due to the division of the complex.

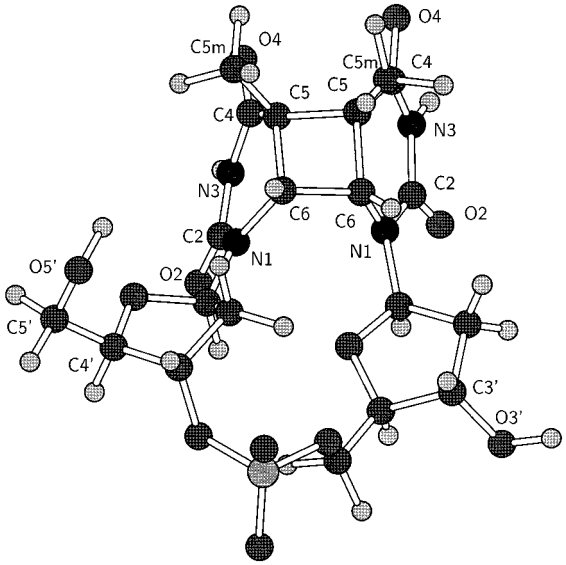

Figure 6. Geometric structure of the acceptor complex.

molecular orbital (HOMO) of the cofactor FAD is well removed from the other filled orbitals (≈1.25 eV) and is a flavin orbital. Four of the seven lowest unoccupied molecular orbitals (MO's) are also localized on the flavin; the other three lowest unoccupied MO's are adenine orbitals. Since a transition from a state on the flavin to one of the latter states upon excitation is less probable than a transition to excited flavin orbitals, a state that is localized on the flavin is considered to be the initial state of the electron-transfer reaction. The adenine states are close in energy to the states on the flavin and, as will be seen below, serve as bridging intermediates in the process of electron transfer.

In one-electron approximation, the donor state is represented by the LUMO of the flavin (144) (compare Figure 5). The LUMO, together with the next lowest unoccupied orbitals and the HOMO, is a π-orbital. It has its largest amplitudes at the C10 atom and the C4═O4 carbonyl group. That is, the main amplitude of the LUMO resides on the isoalloxazine ring.

The LUMO + 1 of the flavin (147) has its main amplitude on the carbonyl groups C2═O2 and C4═O4 and is mainly

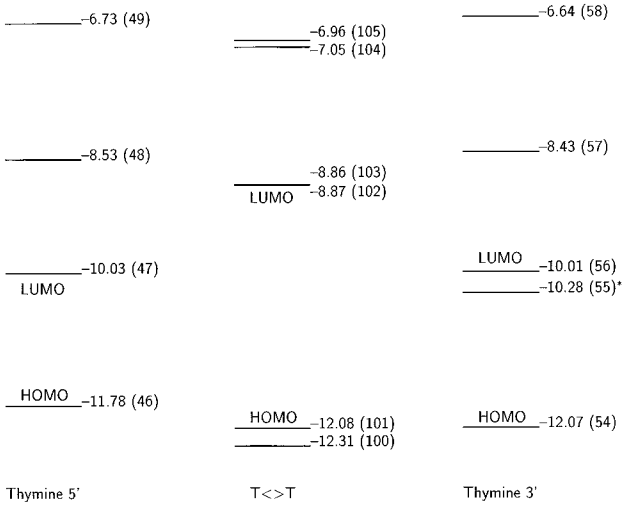

Figure 7. Electronic structure of the thymine dimer in the extended Hückel approximation for the structure used in docking with correlation diagram of the highest occupied and lowest unoccupied molecular orbitals of the dimer with those of the 5′ and the 3′ thymines. The 55 state on the 3′ thymine has 99% of its amplitude on the phosphate.

localized on the isoalloxazine ring, too. The next two unoccupied molecular orbitals (148 and 149) reside on the benzene ring.

3.4. Electronic Structure of the Thymine Dimer. The initial geometry of the thymine dimer is taken from the structure obtained with a molecular dynamics simulation on a segment of damaged DNA duplex[26] (Figure 6). The two highest occupied (100 and 101) and four lowest unoccupied MO's (102−105) are localized on the thymine bases. The four lowest unoccupied orbitals are grouped in two pairs of closely spaced levels (Figure 7). The doublets are due to mixing of corresponding π-orbitals on the individual thymine rings. The LUMO and LUMO + 1 of the dimer (102 and 103) are largely combinations of the 48 and 57 orbitals on each entity. The LUMO's of the individual rings (47 and 56) are not present in the frontier orbitals of the dimer, since they are involved in the formation of the covalent bonds within the cyclobutane ring. The LUMO of the dimer (102) has its main amplitude at the carbonyl groups, C2═O2 and C4═O4. All other atoms have small contribution to the wave function. The LUMO + 1 has a similar character.

3.5. Estimate of Electronic Coupling from Experimental Data. According to the theory of nonadiabatic electron transfer, the rate is given by[51]

$$k = \frac{2\pi}{\hbar}|T_{DA}|^2\rho_{FC} \tag{6}$$

The Franck−Condon factor in the classical approximation is

$$\rho_{FC} = \sqrt{\frac{1}{4\pi\lambda k_B T}}\exp\left(-\frac{(\lambda + \Delta G_0)^2}{4\lambda k_B T}\right) \tag{7}$$

where λ is the reorganization energy and ΔG_0 is the reaction free energy. Typical values for λ in proteins are $\lambda = 1-0.5$ eV. To estimate an upper limit of ρ_{FC}, we set the exponential factor equal to 1 (activationless transfer, maximum rate). At $T = 275$ K we get

$$\rho_{FC} \approx (2.3 - 3.2) \times 10^{-4}\ \mathrm{cm}^{-1} \tag{8}$$

For the corresponding rate for DNA photolyase and a thymine dimer[14] of (100 ps)$^{-1}$ we expect a coupling strength of at least

(51) Marcus, R. A.; Sutin, N. *Biochim. Biophys. Acta* **1985**, *811*, 265.

$$|T_{DA}| \approx 5-6 \text{ cm}^{-1} \quad (9)$$

Together with the rate[14] $(400 \text{ ps})^{-1}$ at $T = 90$ *K*, the exponential factor in ρ_{FC} is calculated to be $\exp(-260\ K/T)$. Thus, the expected coupling strength is about 10 wavenumbers or larger.

3.6. Transfer Matrix Element. To calculate the transfer matrix element, either by the method of avoided crossing or by the method of tunneling currents, the Hamiltonian of the entire system has to be diagonalized. The complex of the whole protein and a thymine dimer contains 7636 atoms with 19 312 valence orbitals. To obtain a system that is accessible to direct diagonalization and retain the part that is important for electron transfer, we substitute the protein by its active site. If the polypeptide chain is broken, the N and C terminals are capped with H and OH, respectively, using the InsightII program.[34] A truncation radius of 6 Å results in a complex with 44 residues (including donor and acceptor complexes), containing 850 atoms with 2224 valence orbitals. The structure obtained in that way was also used in the MD simulations.

A number of calculations of electronic coupling were performed on different structures obtained in the docking procedure, using different states of FAD and dimer as donor and acceptor states. The computed values of the transfer matrix element were in the range of 1 to 10 cm^{-1}. As expected, there is a strong dependence of electronic coupling upon the geometry of the structure and upon the donor and acceptor states chosen for calculations. The important conclusion from this series of calculations is that the dimer has to be in a close contact with the FAD in order to obtain the coupling strength consistent with experimental data. For intermolecular distances of 5 Å or larger we could not obtain electronic coupling that is larger than 1 cm^{-1}, while the experimental value is at least 10 cm^{-1}.

To explore the influence of the donor wave function on T_{DA}, beside the LUMO the three next lowest unoccupied molecular orbitals of flavin in the initial configuration were considered as donor states, too.

The major finding is that the donor states with maximal amplitude on the benzene ring of the flavin (LUMO + 2 and LUMO + 3) do not have a stronger coupling than the LUMO and LUMO + 1, which are localized on the isoalloxazine moiety.

The fact that the coupling strength for the wave functions located on the benzene ring, which is displaced 4.9 Å from the isoalloxazine moiety mainly into the direction of the acceptor, is not larger than the coupling of the LUMO and LUMO + 1 localized on the isoalloxazine is an indication that the electron transfer between the flavin and the dimer is not direct.

For the calculation of the matrix element along the dynamics trajectory the method of tunneling currents was used. Sensitivity analysis (see below) shows that the protein matrix does not participate in the electron transfer process. Accordingly, only the FAD, the dimer, and adjacent water molecules were left in the structures used for the calculation of matrix element.

The LUMO of the flavin molecule was chosen as the donor state in these calculations, while both LUMO and LUMO + 1 of the dimer were chosen as the acceptor states. The graph of the matrix element along the dynamics trajectory of T<>T dimer with the LUMO of the dimer as the acceptor state is shown in Figure 8. The initial time refers to the beginning of the main molecular dynamics run. The coupling matrix element experiences fluctuations with an amplitude in the range of 30 cm^{-1} (maximal value of the matrix element computed with LUMO of the dimer as the acceptor state is about 33 cm^{-1}). A qualitatively similar behavior was found for all four dimers T<>T, T<>U, U<>T, and U<>U.

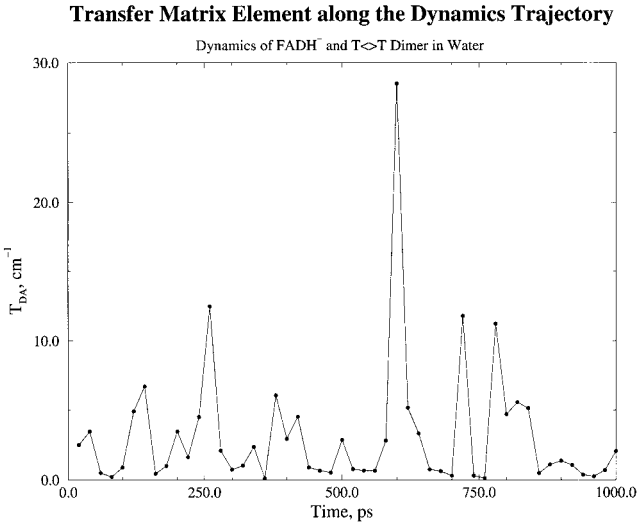

Figure 8. Transfer matrix element along the dynamics trajectory: rms matrix element, 6 cm^{-1}; variance of matrix element, 5 cm^{-1}.

The rms matrix element for the T<>T dimer was calculated to be equal to 6 cm^{-1} with variance of about 5 cm^{-1}. For other dimers these values are as follows: for the T<>U dimer, the rms matrix element is 7 cm^{-1} with variance of about 5 cm^{-1}; for U<>T, 7 and 5 cm^{-1}; for U<>U, 7 and 6 cm^{-1}. For all four different dimers the rms matrix element calculated with LUMO + 1 of the dimer as the acceptor state was somewhat smaller—from 3 cm^{-1} (U<>T) to 5 cm^{-1} (U<>U).

The values of electronic coupling obtained for the LUMO on the dimer are consistent with experimentally determined rates of transfer. The absence of the pronounced differences in the values of matrix elements for different dimers is also consistent with experimental results. Since the variance of the matrix element is of the order of the rms value of the matrix element itself, it is impossible to reliably predict the difference in the averaged matrix elements as small as a factor of 1.7, which corresponds to the observed 3-fold difference in the rate.

To check the possible influence of the protein itself on the coupling between FAD and the dimer, calculations were performed with and without protein. When the protein matrix is removed from the structure found in the docking, the coupling matrix element changes only by 10−30% compared with that of the whole system. The fact that this approximation does not reduce the electronic coupling strength underlines that the polypeptide chain does not play a role in the electron-transfer event.

While the long-range features of the coupling show an exponential decay with increasing intermolecular distance, for distances between the flavin and the dimer below 3 Å the coupling strength is seemingly uncorrelated with the shortest closest distance, r_{df}. In particular, for the configuration with the strongest coupling obtained in the docking procedure the r_{df} value is 2.8 Å, which is nearly 0.5 Å larger than that of van der Waals contact. This leads to conclusion that the shortest distance between the flavin part of FAD and the dimer, r_{df}, does not have much influence on the electronic coupling once the dimer is deep in the binding site.

On the other hand, the dependence of matrix element upon the distance between the dimer's C4═O4 groups, where the acceptor wave function is mainly localized, and the adenine part of FAD, r_{da}, is pronounced. This dependence indicates the involvement of the adenine in electronic coupling. In contrast to the flavin at the bottom of the binding site the adenine is located closer to the exterior of the protein. Furthermore, the

adenine atom closest to the C4═O4 groups is always one of the two hydrogens at the AN6 nitrogen atom, allowing the formation of a hydrogen bond between the dimer and the adenine.

Thus, we conclude that the adenine has the role of a bridging intermediate in the electron-transfer reaction. Sensitivity analysis of the matrix element confirms this conclusion. The removal of the adenine from the structure obtained in docking leads to a pronounced drop of electronic coupling.

Both the influence of the distance between C4═O4 carbonyl groups of dimers and AN6 atom of the adenine and the sensitivity of the matrix element to the removal of the adenine indicate that the adenine has the role of the bridging intermediate in the electron-transfer reaction. An analysis of the eigenstates of the configurations along the dynamic trajectories also confirms this conclusion. The energy of the tunneling electron is close to that of the lowest unoccupied adenine orbital. This energy varies along the dynamic trajectory. We noticed that the peaks of the matrix element occur when the energy of the tunneling electron is particularly close to that of one of the adenine states. The proximity of energies of the transferring electron and the states on the adenine, and the position of the adenine with respect to both the dimer and the flavin, leads to an enhanced interaction between the donor and acceptor in the repair reaction.

The final remark in this section is on the role of the solvent water molecules in the electron-transfer process in this system. There seems to be no correlation between the average value of the matrix element and the number of water molecules located in the catalytic site. And there are no water molecules right between the dimer and the FAD. This leads to the conclusion that water does not play a significant direct role in electronic coupling. However, it does not exclude the possibility that water may play a role in overall electron transfer by modifying the driving force and the reorganization energy of the reaction. Also water seems to significantly affect the position of the dimer in the bining pocket. We performed a molecular dynamics simulation of the system in the absence of solvent water. In the configurations obtained in that simulation all dimers, except for T<>T, are on average significantly farther from the FAD than in simulations with water. Accordingly, the averaged transfer matrix element for these structures was of the order of $1-2$ cm^{-1} and much different from that of the T<>T pair. This is in worse agreement with experimental data than the results obtained with water.

4. Summary and Conclusions

In the study of the mechanism of DNA photorepair by photolyases, the key questions that need to be answered are those about the structure of the DNA/photolyase complex and the mechanism of electron transfer from the excited $FADH^-$ to the T<>T dimer of the damaged DNA. Using the DOCK 4.0 program,[22,23] we explored possible configurations of the dimer within the active site of *E. coli* photolyase[5] and used molecular dynamics simulation, which included solvent water, to study the dynamic behavior of the system. For various positions of the dimer within the catalytic site along the dynamic trajectory, we calculated electronic coupling between the lowest unoccupied molecular orbitals of the flavin and the dimer. The electronic structure of the system was treated within the extended Hückel approximation.

In principle, electron transfer can occur over long distances.[52,53] The magnitude of electronic coupling and the rate of electron-transfer reactions, however, decrease exponentially with the distance between donor and acceptor complexes. We find that, for the matrix element to be consistent with the experimentally determined rate of electron transfer,[14,16,17] the dimer must occupy a specific position deep inside of the cavity leading to $FADH^-$. In the T<>T configuration with the largest coupling strength, the methyl group of the thymine 3′ of the dimer closely approaches the C8m methyl group of the flavin. In this configuration, the C4═O4 carbonyl groups of the base pair are in close proximity of the NH_2 of the adenine.

Analysis of the dynamic behavior of the dimer in the binding pocket shows that the critical parameter governing the strength of the coupling with FAD is not the distance between the dimer and the edge of the flavin molecule exposed to solvent. The distance between the C4═O4 carbonyl groups of the dimer and the adenine part of FAD is more important. This and other findings lead us to conclude that the primary electron transfer between $FADH^-$ in DNA photolyase and a thymine dimer bound in its active site is indirect via the adenine moiety acting as an intermediate.

The proposed electron-transfer mechanism utilizes the unusual conformation of the $FADH^-$ cofactor specific for photolyases, in which the isoalloxazine ring and the adenine are in close proximity;[5,6,24] this proximity leads to a strong overlap between their π-systems. Moreover, the dimer is bound in the pocket in such a way that its LUMO, which has the largest amplitude on the C4═O4 carbonyl groups, overlaps with the states on the adenine, too. The strong coupling of the donor and acceptor states with the same intermediate electronic states of the adenine causes an effective superexchange coupling between donor and acceptor, without the necessity of their direct overlap. This type of coupling makes an effective electron transfer to pyrimidine dimers possible, while protecting the flavin from being oxidized by other agents. The main role of the protein is to adjust the dimer properly with respect to its catalytic cofactor.

An immediate consequence of this model is that a mutation of FAD that destroys the aromaticity of the adenine without affecting the dimer and cofactor binding properties of photolyase should affect the rate and the quantum yield of the repair. The essential role of the C4═O4 carbonyl groups of the thymine dimer in establishing the effective coupling is reflected in the about 20-fold lower quantum yield of repair for C<>C than for T<>T dimers[13] supposed that both types of pyrimidine dimers are bound in a similar way by photolyase.

The global configuration of the thymine dimer in the active site of photolyase that we found in docking and dynamic simulations has the correct order of magnitude of the electronic coupling between the dimer and FAD. This configuration qualitatively agrees with the one discussed on the basis of the active site structure by Park et al.[5] If the position determined is indeed correct, then the principal question is how the thymine dimer gets into the binding site in the configuration described, or close to it, when it is part of DNA. Similarity between positions of a dinucleotide dimer and a T<>T dimer as a part of DNA is indicated by experiments on quantum yield of repair[13] with similar values in both cases. A recent mutagenesis study in connection with a modeling assay[25] provides further experimental evidence for this binding mode, which requires a significant extra deformation of the damaged strand in the vicinity of the lesion.[54]

The difficulty to model such a singular process as the dimer flipping out of DNA helix has been recently demonstrated by

(52) Moser, C. C.; Keske, J. M.; Warncke, K.; Farid, R. S.; Dutton, P. L. *Nature* **1992**, *355*, 796.
(53) Scott, R. A. *JBIC* **1997**, *2*, 372.
(54) Kunkel, T. A.; Wilson, S. H. *Nature* **1996**, *384*, 25.

two molecular dynamics simulation studies.[20,21] The configurations in which the dimer is in close contact with the FAD cofactor were not found in these simulations. This can be explained by many factors; for example, it could be due to the choice of the initial conditions and the duration of the simulation runs which do not exhaust the entire configurational space. It might be also due to a principle limitation of the method in that the harmonic force fields used, in the simulations, inadequately disfavor the bond stretching events necessary for the dimer flipping reaction. That the dimer does not completely enter the binding site is difficult to believe, given the high rate of electron transfer and the strong electronic coupling between the dimer and FAD. Our molecular dynamics study shows that the configuration we found in docking is stable and the dimer does not get out of it. Importantly, it has the correct order of magnitude of the coupling, which is a very sensitive indicator of the position of the dimer in the binding site. Why would nature reserve a space in the enzyme and then not allow the only occurring substrate to occupy it, and why would the electron-transfer reaction occur in a suboptimal way? If our conclusions are correct, one of the critical issues in further study of the DNA repair mechanism by photolyase will be to understand the mechanism of dimer flipping out of DNA.

Acknowledgment. We are grateful to Dr. Iraj Daizadeh for a series of stimulating discussions, assistance with computer programs, and for indicating a number of valuable references; we are also grateful to Dr. Jian-xin Guo for help with molecular dynamics simulations and ab initio quantum chemical calculations. Mike Bailey of the San Diego Supercomputer Center's TeleManufacturing Facility has produced 3-D solid models of the protein and a piece of damaged DNA used in this study and is gratefully acknowledged. We also acknowledge technical assistance provided by dock@francisco.ucsf.edu. This work was supported by the research grant from the National Institutes of Health (GM54052-02) and by the fellowships from the Sloan and Beckman Foundations.

JA993784T