Published in final edited form as:
J Am Chem Soc. 2018 February 28; 140(8): 2853–2861. doi:10.1021/jacs.7b11926.

Determinants of Photolyase's DNA Repair Mechanism in Mesophiles and Extremophiles

Benjamin J. G. Rousseau[†], **Shoresh Shafei**[†], **Agostino Migliore**[*,†], **Robert J. Stanley**[‡], **David N. Beratan**[*,†,‖,§]

[†]Department of Chemistry, Duke University, Durham, North Carolina 27708, United States

[‖]Department of Physics, Duke University, Durham, North Carolina 27708, United States

[§]Department of Biochemistry, Duke University, Durham, North Carolina 27710, United States

[‡]Department of Chemistry, Temple University, Philadelphia, Pennsylvania 19122, United States

Abstract

Light-driven DNA repair by extremophilic photolyases is of tremendous importance for understanding the early development of life on Earth. The mechanism for flavin adenine dinucleotide repair of DNA lesions is the subject of debate and has been studied mainly in mesophilic species. In particular, the role of adenine in the repair process is poorly understood. Using molecular docking, molecular dynamics simulations, electronic structure calculations, and electron tunneling pathways analysis, we examined adenine's role in DNA repair in four photolyases that thrive at different temperatures. Our results indicate that the contribution of adenine to the electronic coupling between the flavin and the cyclobutane pyrimidine dimer lesion to be repaired is significant in three (one mesophilic and two extremophilic) of the four enzymes studied. Our analysis suggests that thermophilic and hyperthermophilic photolyases have evolved structurally to preserve the functional position (and thus the catalytic function) of adenine at their high temperatures of operation. Water molecules can compete with adenine in establishing the strongest coupling pathway for the electron transfer repair process, but the adenine contribution remains substantial. The present study also reconciles prior seemingly contradictory conclusions on the role of adenine in mesophile electron transfer repair reactions, showing how adenine-mediated superexchange is conformationally gated.

Graphical Abstract

[*]**Corresponding Authors**: agostino.migliore@duke.edu, david.beratan@duke.edu.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.7b11926.

Structure of the *A. nidulans–CPD* complex (PDB)

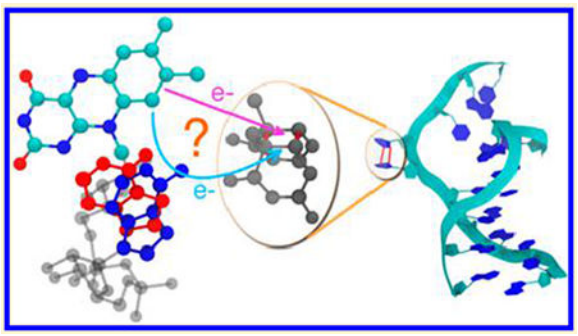
Structure of the *E. coli–CPD* complex (PDB)

Structure of the *S. tokodaii–CPD* complex (PDB)

Structure of the *T. thermophilus–CPD* complex (PDB)

MD parameters and simulations; system RMSDs over the MD production runs; details of the flavin geometry; and pathway analysis, including the distribution of the initial and final electronic states on the individual atoms and the effect of water on the extent of adenine ET mediation role (PDF)

The authors declare no competing financial interest.



INTRODUCTION

Early life on Earth developed before the advent of oxygen-evolving plants,[1] encountering extremes of temperature and intense ultraviolet (UV) irradiation. UV irradiation causes photochemical damage to adjacent pyrimidine bases in DNA, forming a cyclobutane pyrimidine dimer (CPD) moiety. This damage can induce mutagenesis and finally lead to cell death.[2] Photolyases repair UV-induced lesions in DNA using photoinduced electron transfer (ET). Understanding the DNA-repairing ET mechanism of mesophiles and extremophiles may help to shed light on the early evolution of life on Earth. Furthermore, the photoinduced ET function of DNA photolyases makes their structure of great interest in the context of protein engineering and design.[3,4]

The repair mechanism of DNA photolyases has been studied extensively.[5–12] DNA repair by photolyase is initiated when an antenna chromophore (Figure 1) absorbs near UV/blue light and transfers its excitation energy to flavin adenine dinucleotide ($FADH^-$).[13] In the photolyase–DNA complex, the pyrimidine dimer flips out of the DNA backbone and comes into van der Waals contact with the excited $FADH^-$, which is accessible to DNA through a gap in the protein helical domain.[14–18] The excited $FADH^-$ can then donate an electron to the CPD (forward ET), which leads to its cleavage.[14,17,19–21] After DNA repair, back ET restores the negative charge of the FADH cofactor.[22]

The detailed mechanism of the critical forward ET step in the enzyme–CPD complex is poorly understood. Depending on the adenine orientation relative to the flavin and CPD, the electron can tunnel from the flavin donor to the CPD acceptor without the assistance of virtual electronic states localized on the adenine moiety.[23] A previous theoretical study[24] identified a shortest tunneling pathway that starts from the terminal methyl group of the flavin and ends on the thymine at the 3′ side of CPD. Portions of the protein cannot enter the gap between the FADH cofactor and the CPD (confirmed by inspection of the molecular dynamics trajectories produced in this study), while water molecules can assist the tunneling process. We refer to the ET repair reaction without adenine mediation as direct ET; the direct mechanism includes through-space and water-mediated tunneling.

The “U” shape of $FADH^-$ in DNA photolyases (Figure 2) can allow the adenine to come into proximity with CPD, thus enabling adenine-mediated electron superexchange[25] or hopping from flavin to adenine to CPD.[24] The electron hopping mechanism was excluded in earlier

studies.[25–27] Experiments by MacFarlane and Stanley,[8] as well as later experiments by Kao et al.,[27] found no ET from the flavin to adenine in the absence of DNA substrate, thus ruling out the hopping mechanism. A recent theoretical study by Lee et al.[24] supports the possible coexistence of charge hopping and tunneling (either mediated by adenine or not). Wang et al. recently suggested a proton-coupled ET mechanism that proceeds via a water wire.[28]

Studies of the forward ET in photolyases provide evidence of both direct ET and adenine-mediated superexchange. Stuchebrukhov et al., in their computational analysis, found that the forward ET rate was sensitive to the presence of adenine and to its distance from CPD.[25,26] Their theoretical finding was corroborated by recent experiments of Liu et al. using ultrafast absorption spectroscopy.[29,30] The experiments were interpreted in the context of non-adiabatic ET theory, ruling out two-step hopping through the adenine and supporting a superexchange mechanism.[30] In fact, it was argued that, if hopping were feasible, the first ET step from flavin to adenine could occur in the absence of the DNA substrate, but Kim et al.,[13] MacFarlane and Stanley,[8] and Liu et al.[29,30] did not find rapid decay of the $FADH^-$ excited state (caused by ET to the adenine) when the DNA substrate was absent.[30] However, the change in local environment that is produced by the DNA presence may influence, both geometrically and energetically, a potential ET step from the flavin to the adenine. Furthermore, recent experiments, using photolyases with adenine-modified FAD,[31] support the possible occurrence of two-step hopping, but a radical anion intermediate of the modified adenine was not observed. Electronic structure calculations by Weber et al. found that the protein environment surrounding the FADH anion favors strong electronic overlap between the flavin radical and the adenine moiety[32] (and therefore a large electronic coupling between them[33]). The crystal structure of a DNA lesion bound to photolyase after *in situ* repair shows that the adenine bridges the flavin and CPD with two H bonds, thus supporting adenine-mediated superexchange.[20]

The theoretical and experimental studies mentioned above support a one-step tunneling mechanism for ET rather than hopping, but these studies do not quantify the relative importance of direct ET vs adenine-mediated superexchange. Theoretical analysis by Prytkova et al.[23] found that the lowest electronic excited state of $FADH^-$ is localized on the side of the flavin farthest from adenine and closest to CPD (“proximal side”), suggesting a role for direct ET.

The recent theoretical study of Lee et al.[24] identified eight possible pathways for ET from the excited $FADH^-$ to CPD. Two of these pathways involve sequential electron hopping through adenine. The other six ET paths involve tunneling, and start from the two lowest-lying singlet excited states of $FADH^-$, which originate in localized $\pi\pi^*$ excitations of the flavin. The computations[24] did not distinguish between direct and adenine-mediated ET pathways.

Here, we focused on the mechanisms of electron tunneling from flavin’s excited electronic states to CPD. We used quantum chemistry calculations, molecular dynamics (MD) simulations, and electron tunneling pathway analysis to assess the role of adenine as a superexchange mediator in four photolyases from mesophiles and extremophiles. We found that adenine conformational flexibility can cause the charge transfer repair reaction to occur

either by direct ET or by adenine-mediated superexchange, and we explored how the balance of the two ET mechanisms varies among proteins that operate at different temperatures.

MATERIALS AND METHODS

Enzymes Studied.

We studied two mesophilic (*Anacystis nidulans*[34] and *Escherichia coli*[15]), one thermophilic (*Thermus thermophilus*),[19] and one hyperthermophilic (*Sulfolobus tokodaii*)[35] photolyases (Table 1). The structure of *S. tokodaii* photolyase is more similar to that of *A. nidulans* photolyase (sequence similarity of about 35%) than to the structures of *E. coli* (32%) or *T. thermophilus* (28%) photolyases. However, all four proteins have large structural homology in the $FADH^-$ binding domain.

Molecular Docking.

We used the structure of *E. coli* photolyase[15] in a complex with the CPD moiety from ref 25. For *A. nidulans*, *T. thermophilus*, and *S. tokodaii* photolyases, we docked the CPD to photolyase using the Autodock Vina program.[37]

The docking was preceded by thermal equilibration of the separated enzyme and CPD. Since the crystallization temperatures do not match the bacterial physiological growth temperatures (at which the photolyases operate), classical MD simulations were used to equilibrate the DNA photolyases at their physiological temperatures (we carried out 10^5 energy minimization steps, followed by 50 ps of equilibration at constant temperature, 200 ps of equilibration at constant temperature and pressure, and 2.5 ns of MD simulation). The CPD moiety was equilibrated at the physiological temperatures of the different photolyases using a similar protocol (see Supporting Information (SI)). Five structure snapshots separated by 0.5 ns were taken from the final MD trajectory of each protein and the CPD moiety at the pertinent temperature. The two molecular components were then docked, producing 25 conformations for each complex. The initial conformations of the *T. thermophilus* and *A. nidulans* complexes for the ET study were selected, from the 25 conformations available for each of them, as the docked structures that were most consistent with structural information on protein–DNA complexes obtained in previous crystallographic studies of these two enzymes.[19,20] By exploiting the homology of different photolyases in their active domains[15,19,35] (this homology is strictly related to the highly conserved CPD recognition mechanism in DNA photolyases[20,25]), we selected the docked structure of the $FADH^-$–CPD complex in *S. tokodaii* that has the minimum geometric difference from the active domain structure of *A. nidulans* photolyase (PDB: 1QNF[34]). The initial structures chosen for the four photolyases are shown in Figure 3.

MD Simulations.

The Amber12 software package[38] and AMBER force fields (ff12SB) were used to describe the interactions. Parametrizations for CPD and $FADH^-$ were taken from a modified version[39] of the force field developed by Antony et al.[25] The overall system was charge neutralized with Na^+ ions[40] and solvated in a box of TIP3P water extending 10.0 Å on each

side of the complex.[41] The SETTLE algorithm[42] was used to constrain the hydrogen atom positions relative to the heavy atoms. The particle mesh Ewald method[43] was used to calculate the electrostatic interactions in the system. A Langevin thermostat and piston[44,45] were used to control the temperature and pressure (see SI). Trajectories of 60 ns each were produced for the four systems using the NAMD software package.[46] A total of 10 000 conformations were extracted from the last 50 ns of the MD for ET analysis. The ET pathway analysis was performed over time ranges of different length (25 and 50 ns) to test the robustness of the results with respect to the simulation time.

Tunneling Pathways Analysis.

The pathway model[47–50] identifies through-bond and through-space interactions that mediate tunneling between the donor (D) and acceptor (A).[51,52] The D–A pathway-mediated coupling (H_{DA}) is approximated as

$$H_{DA} \propto \prod_i \varepsilon_C(i) \prod_j \varepsilon_H(j) \prod_k \varepsilon_S(k) \tag{1}$$

In eq 1, *i, j,* and *k* denote the through-covalent bond, through-hydrogen bond, and through-space steps between D and A. ε_C, ε_H and ε_S are the decay factors for the corresponding superexchange steps:[53]

$$\begin{aligned} \varepsilon_C &= 0.6 \\ \varepsilon_H &= \varepsilon_C{}^2 e^{-\beta_S(R_H - 2.8)} \\ \varepsilon_S &= \varepsilon_C e^{-\beta_S(R_S - 1.4)} \end{aligned} \tag{2}$$

where R_H (R_S) is the length in angstroms of the H-bond (through-space) step and β_S = 1.1 Å^{-1}.[53,54] The strongest electron tunneling pathway is produced by the combination of steps that maximizes the product in eq 1.

We used the VMD[55] pathways plugin[53] to identify and to visualize the strongest ET pathways from $FADH^-$ to CPD. For each photolyase–DNA complex, pathway analysis was used on all selected MD snapshots to compute the fraction of strongest coupling pathways that contain adenine atoms. This fraction is a measure of adenine's role in mediating D-to-A superexchange.

Electron Donor and Acceptor.

We carried out density functional theory (DFT) calculations on the pruned flavin (Figure 4) to describe the donor charge distribution in the pathway model, rather than to assume a uniform charge distribution. This enables improved assessment of adenine's role as a superexchange mediator. Two facts allow us to limit the DFT analysis to the pruned flavin moiety (without including the adenine in the excited electronic state analysis): (i) We are treating electron tunneling from the excited flavin donor, so we can disregard the charge-transfer excited states that correspond to ET from flavin to the adenine moiety, which were considered to study the hopping mechanism in ref 24. (ii) D-to-A tunneling proceeds from excited electronic states that are localized with the flavin.[24] The excitations were identified and characterized using the NWChem[56] and Gaussian[57] software packages, using time-

dependent DFT (TDDFT) at different levels of computational accuracy (basis set: cc-pVTZ;[58] exchange-correlation functionals: PBE0,[59] M06-2X,[60] CAM-B3LYP,[61] and B3LYP[62] in conjunction with the Casida-Salahub'00 asymptotic correction[63]). These flavin excitations consist essentially of an electron transition from the Kohn–Sham highest occupied molecular orbital (HOMO) of $FADH^-$ to its lowest unoccupied molecular orbital (LUMO) or to the LUMO+1 (see next section).

For each of the two flavin excited electronic states considered (that is, in the single-particle approximation, for the LUMO and the LUMO +1), we computed the strongest coupling pathways starting from all atoms with non-zero amplitude of the pertinent molecular orbital. The pathways were weighed by the contributions of the starting and ending atoms to the molecular orbital, which were calculated using the Mulliken population analysis.[64] The wave function ψ (either the LUMO or the LUMO+1) is expanded in a non-orthogonal set of normalized basis functions $\{\chi_i\}$:

$$\psi = \sum_{i=1}^{N} c_i \chi_i \tag{3}$$

The integrated electron density associated with ψ is

$$\| \psi \| \equiv \int |\psi(\mathbf{r})|^2 d\mathbf{r} = \sum_p \mu_p^{\mathrm{D}} \tag{4}$$

where

$$\mu_p^{\mathrm{D}} \equiv \left[\int |\psi(\mathbf{r})|^2 d\mathbf{r} \right]_p = \sum_{i \in p} \sum_{j=1}^{N} c_i c_j S_{ij} \tag{5}$$

is a measure of the fraction of $\|\psi\|$ on atom p. S_{ij} is the overlap of χ_i and χ_j. In the sum, p indicates the set of electron basis functions localized on the given atom. μ_p^{D} is an atomic electron-localization factor,[65] identified with the Mulliken electron population of atom p when the molecular orbital ψ is occupied by the excited electron charge.

The singly occupied molecular orbital (SOMO) of the anionic CPD was used in the pathway analysis to describe the distribution of transferred charge on the acceptor. The SOMO was calculated at the M06-2X/cc-pVTZ level of theory, and was similarly partitioned among the CPD atoms, thus generating a further atomic weighting factor μ_q^{A} for the ET pathways terminating on atom q of the acceptor. An ET pathway connecting atom p of D to atom q of A is therefore weighted by the factor $\mu_p^{\mathrm{D}} \mu_q^{\mathrm{A}}$.

The LUMO and LUMO+1 of the flavin, as well as the SOMO of the CPD moiety, depend weakly on the nuclear geometry,[66] and this dependence is neglected.

Role of Adenine in Forward ET.

The ET pathways were used in the following protocol to assess the contributions of the direct ET mechanism and the adenine-mediated superexchange to the forward ET:

1. For each MD snapshot, the pathways plugin[53] was used to identify the strongest coupling pathway linking each D–A atom pair (p, q).
2. For each (p, q) pair, a Python program was used to calculate the fraction f_{pq}^{DA} of system snapshots with adenine atoms in the strongest pathway.
3. The average fraction f of strongest ET pathways that include adenine (that is, our measure of adenine-mediated superexchange contribution to forward ET) was calculated as[67]

$$f = \sum_{p,\,q} \boldsymbol{\mu}_p^{\mathrm{D}} \boldsymbol{\mu}_q^{\mathrm{A}} f_{pq}^{\mathrm{DA}} \tag{6}$$

This fraction f ranges between 0 and 1. Zero corresponds to direct ET. One corresponds to the opposite limit in which $f_{pq}^{\mathrm{DA}} = 1$ for any p,q. That is, when $f = 1$, at least one adenine atom is included in all of the strongest ET pathways connecting each D–A atomic pair in the different snapshots (entirely adenine-mediated tunneling). Adenine-mediated superexchange is at play in a given structural snapshot even if adenine atoms are involved in only a fraction of the strongest ET pathways between D–A atomic pairs. However, f is less than one in this case. In addition, there can be structure snapshots where adenine is not involved in the strongest ET pathway but participates in a pathway of very similar strength (multiple tunneling pathways and their possible interferences were not considered in this analysis). Thus, a value of f that is an appreciable fraction of unity indicates an important superexchange mediation role for adenine in repair ET.

RESULTS AND DISCUSSION

Flavin and CPD Electronic States.

Our TDDFT analysis (Table 2) finds that the two lowest-lying singlet excited electronic states of flavin, S1 and S2, correspond approximately to electron transitions from the HOMO to the LUMO and to the LUMO+1, respectively. Only calculations that use the CAM-B3LYP exchange-correlation functional describe the S2 state as arising from an electron excitation to the LUMO+2. This result arises from the interchange of the LUMO+1 and LUMO+2 that can be ascribed[68] to the incompleteness of the basis set and the small energy difference (less than 0.1 eV) between the two molecular orbitals. B3LYP and PBE0 functionals underestimate the S2 excitation energy compared to the observed absorption maximum of $FADH^-$ in DNA photolyase at 3.5 eV.[69] The two other functionals, with a larger Hartree–Fock exchange component, produce an excitation energy for S2 that is only slightly larger than the experimental value. Overall, the computed S1 and S2 state energies are in agreement with the experimental values, which correspond approximately to wavelengths of 425 and 360 nm, respectively.[69,70] In fact, considering the averages of the excitation energy values from the different functionals as our best estimates and using confidence intervals of 95%, we find the excitation wavelengths (413 ± 12) nm and (364 ± 11) nm for S1 and S2, respectively. Previous experimental and theoretical studies (see ref 24 and references therein, as well as ref 70) describe the S1 excitation energy as lying between

the computed PBE0 and M06-2X values, while only the latter functional produces the observed order of magnitude for the small oscillator strength associated with this excitation. Overall, the M06-2X density functional showed the best agreement with the experiment on the properties of the S1 and S2 states, and was thus used to calculate the LUMO and LUMO +1 of the flavin (Figure 5), as well as the SOMO of CPD.

The charge distributions of the LUMO and LUMO+1 on the flavin atoms are reported in detail in SI, Table S1 (which also contains the atomic decomposition of the SOMO of CPD) and shown in Figure 5 (cf. Figures 4 and 5). The C7 and C9A atoms make the largest contributions to the LUMO, in agreement with the results of Prytkova et al.[23] that found significant electronic localization on the proximal side of flavin, especially on C6, C7, C9 and C9A. Our results for CPD are consistent with the SOMO localization on the C2═O2, C4═O4 carbonyl groups reported by Antony et al.[25] However, Table S1 shows that many other flavin and CPD atoms need to be included in the analysis to describe the role of adenine in the forward ET. In fact, D and A atoms with small electron-localization factors can be linked by efficient coupling pathways. In our pathways analysis, we included all possible D–A atom pairs.

Tunneling Mediated by Adenine.

By substituting the localization factors of Table S1 and the f_{pq}^{DA} values calculated from the MD simulations in eq 6, we computed the average fraction f of adenine-mediated strongest ET pathways for each photolyase species (Table 3).

There is a significant adenine-mediated superexchange contribution to the ET repair reaction when CPD is complexed with the photolyases in *A. nidulans* (mesophile) and in the two extremophiles (*T. thermophilus* and *S. tokodaii*) at their physiological temperatures. In contrast, the predominant ET mechanism in the *E. coli* photolyase at its physiological temperature (37 °C) is direct ET, with only about 3% of the strongest ET pathways mediated by adenine. When *E. coli* photolyase is simulated at the same temperature as the *A. nidulans* photolyase (20 °C), the contribution of the superexchange mechanism to the forward ET also becomes important in the *E. coli* species. This change is ascribed to a more nearly locked adenine orientation between flavin and CPD at lower temperatures (see below).

Effects of Conformation, Temperature, and Water on the Adenine-Mediated ET.

The MD trajectory explains the tunneling mediation role of adenine in *E. coli* photolyase as a function of temperature. In this enzyme, at 310 K, the adenine portion of $FADH^-$ frequently exhibits a “tilted” geometry with respect to CPD (Figure 6). At 293 K, adenine rarely exhibits this tilted conformation, remaining largely in a conformation proximal to the DNA lesion that supports superexchange. Our MD simulations of the four photolyase species indicate that adenine can fluctuate between the tilted and proximal conformations, but the population of the tilted conformation is significantly higher in the mesophilic photolyase of *E. coli* than in the photolyases of the other three species studied at their respective physiological temperatures.

Adenine flipping between two geometries in the simulations helps to explain the origins of the debate in the literature regarding the role of adenine in the CPD repair. For example, Prytkova et al.[23] found that direct ET dominates the ET repair reaction in *E. coli* photolyase, which was ascribed to a geometry of the enzyme that is similar to the average geometry obtained in our MD simulation at 310 K. The looser architecture of the flavin and CPD when the adenine is tilted is expected to reduce the D–A electronic coupling and to increase its sensitivity to geometry, in agreement with the small value of the coherence parameter[72–74] reported in ref 23. A different average position of the adenine with respect to the flavin and CPD may have led to the opposite conclusion, namely that adenine mediates forward ET, in the theoretical study by Antony et al.[25] The tilted adenine that we found at 310 K (that is, 10 K above the temperature used in ref 25) may explain the reduced ET-mediation role of adenine at this temperature, and this result motivates future MD studies to assess the population of this geometry that arises at longer times. Reducing the temperature to 20 °C locks the complex in a conformational subspace that favors adenine-mediated tunneling. We conclude that, in *E. coli* DNA photolyase, adenine-mediated superexchange is conformationally gated, with an occurrence that depends on temperature.

The *f* values computed for the *A. nidulans* mesophile and for the two extremophiles indicate an important role for adenine in mediating ET from the photoexcited flavin to the CPD lesion (in agreement with recent experimental findings on mesophiles[12,30]). We also found that water can cause appreciable fluctuations in the value of *f* on a nanosecond time scale (see SI section 4), but the occurrence, or lack thereof, of significant adenine mediation of the forward ET is only species and temperature dependent. That is, irrespective of the presence of water, in *E. coli* photolyase at 37 °C adenine makes a small contribution to the ET mediation, while in the same enzyme at 20 °C, and in the other photolyases studied, adenine significantly contributes to mediate the ET (cf. the upper panel of Table 3 and SI, Table S3). The decrease in *f* due to water mediation of the forward ET (see SI, section S4) is overestimated by considering only the strongest ET pathway, since ET pathways through adenine and water may have similar donor–acceptor couplings (see SI, Figure S2). Future theoretical studies using multi-pathway analysis will help to quantify the water effects on the extent of adenine's ET mediation with higher accuracy. While the increase in temperature seems to cause adenine displacement and consequent loss of its role as a tunneling bridge in the case of *E. coli*, we found a significant adenine role in both extremophiles, despite their significantly higher temperatures of operation. We found an appreciable correlation of the *f* value with the bacterium growth temperature for the first 25 ns of the MD production run. However, the dependence of f on the species at room temperature is of a size comparable to the *f* increase with enzyme temperature. Furthermore, the correlation between *f* and the enzyme temperature is erased by thermal fluctuations over 50 ns of MD simulation (lower panel in Table 3). The adenine contribution to the strongest ET pathway shows a more evident correlation with the bacterial growth temperature if the effect of water on the electron tunneling by superexchange is excluded (SI, Table S3). This finding (taken with the possible coexistence of through-adenine and through-water ET pathways of similar strength) suggests that a correlation may be found in future theoretical investigations by considering multiple electron tunneling pathways and extending the analysis to examine a wider set of photolyase species.

The recent finding of an optimal growth temperature of 33 °C for the PCC 7942 strain of *S. elongatus*[36] suggests that a similar optimal temperature might be found for the PCC 6301 strain. In this event, it would be useful to study (theoretically and experimentally) whether DNA photolyase from *A. nidulans* above 30 °C experiences a loss of superexchange contribution to the forward ET that is similar to the loss found for the *E. coli* photolyase.

The significant extent of functional ET by superexchange through adenine in the extremophiles (and its structural implications, as derived from our MD analysis) is consistent with Klinman's perspective on functional ET in thermophiles,[75] namely that conformational changes enabled at typical thermophile temperatures allow enzymes to attain a closely packed, rigid, and catalytically efficient active site, at the expense of increased flexibility in other parts of the protein.

CONCLUSIONS

We have combined molecular docking, classical MD simulations, electronic-structure analysis, and pathway analysis to quantify the importance of adenine for mediating electron tunneling from the excited photolyase $FADH^-$ to the DNA CPD lesion to be repaired. Our study extends prior theoretical analysis of DNA photolyases to thermophiles and hyperthermophiles, with the aim of understanding photolyase repair mechanisms of relevance to the early stages of life on Earth. The results reported in Table 3 (taken with the relationship, found in this study, between the adenine-mediated repair mechanism and the structure of the FADH–CPD complex) suggest that the early evolution of DNA photolyases favored those species with a sufficiently rigid active domain structure to support adenine-mediated ET, while mesophiles with looser assemblies of the enzyme active domain could appear at a later stage of the photolyase evolutionary history. Therefore, adenine mediation is also expected to influence the efficiency of the biologically important forward ET. This expectation may be tested in future calculations of the coupling between $FADH^-$ and CPD (aiming to find out whether the direct ET mechanism or the adenine-mediated superexchange mechanism leads to faster repair by ET), as well as through theoretical or experimental adenine-mutation studies.

Our results explain the extent of direct ET (through space or water mediated) vs through-adenine superexchange in the forward ET reaction from the flavin to the DNA lesion. Either mechanism can prevail, depending on thermal access to the corresponding photolyase conformations. Our MD simulations indicate the possible occurrence of two adenine geometries that produce the two different tunneling mechanisms, thus providing a consistent explanation of the divergent conclusions on adenine mediation role in previous studies.[23,25] The population of the two adenine conformations is expected to be enzyme and temperature dependent. At 293 K, we found a somewhat higher adenine-mediated superexchange contribution to the forward ET for *A. nidulans* compared to *E. coli* photolyase. Since the temperatures at which these two enzymes function are expected to differ by a few degrees, it will be useful to investigate the role of adenine in the *A. nidulans* DNA photolyase above 30 °C. This investigation could show whether the ET mediation roles of adenine in *A. nidulans* and in *E. coli*. photolyases decrease at similar temperatures, or if there is a temperature

range where adenine's role is significantly different in the two mesophile enzymes (species-dependent adenine role).

Our simulations and *f*-value analysis indicate that adenine is positioned stably between flavin and CPD in extremophile photolyases. In the *E. coli* mesophile photolyase, adenine tends to flip out of the region between the flavin and CPD at the enzyme operation temperature of 37 °C (see SI, Figure S3). For photolyase from *A. nidulans* mesophile, adenine is predominantly located between the flavin and CPD. However, this enzyme might have a working temperature above 30 °C, similar to the case of the PCC 7942 strain of this bacterium,[36] with possible effects on the stability of adenine's position between the flavin and CPD. Simulations of *A. nidulans* DNA photolyase at temperatures above 30 °C (as well as experimental assessment of the bacterial growth temperature), and of extremophilic DNA photolyases at temperatures different from the optimal growth temperatures, would allow us to establish whether the DNA photolyases of thermophiles and hyperthermophiles, compared to those in mesophiles, maintain the adenine in positions that support superexchange ET over wider temperature ranges.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank Dr. I. Balabin for help with the pathways plugin, Prof. A. A. Stuchebrukhov for providing the structure of *E. coli* photolyase with docked CPD, Prof. T. Yamato for providing $FADH^{-}$ and CPD force fields, and Dr. Tomasz Janowski for technical support. This research was supported by the NASA Exobiology Program (Grant 80NSSC17K0033).

REFERENCES

(1). Blankenship RE; Hartman H Trends Biochem. Sci 1998, 23, 94. [PubMed: 9581499]
(2). Essen LO; Klar T Cell. Mol. Life Sci 2006, 63, 1266. [PubMed: 16699813]
(3). Russell RJM; Taylor GL Curr. Opin. Biotechnol 1995, 6, 370. [PubMed: 7579645]
(4). Vogt G; Argos P Folding Des. 1997, 2, S40.
(5). Sancar GB; Jorns MS; Payne G; Fluke DJ; Rupert CS; Sancar A J. Biol. Chem 1987, 262, 492. [PubMed: 3539941]
(6). Kim ST; Sancar A Photochem. Photobiol 1993, 57, 895. [PubMed: 8337263]
(7). Stanley RJ Antioxid. Redox Signaling 2001, 3, 847.
(8). MacFarlane AW IV; Stanley RJ Biochemistry 2003, 42, 8558. [PubMed: 12859203]
(9). Sancar A Chem. Rev 2003, 103, 2203. [PubMed: 12797829]
(10). Kao YT; Saxena C; Wang L; Sancar A; Zhong D Cell Biochem. Biophys 2007, 48, 32. [PubMed: 17703066]
(11). Weber S Biochim. Biophys. Acta, Bioenerg 2005, 1707, 1.
(12). Liu Z; Wang L; Zhong D Phys. Chem. Chem. Phys 2015, 17, 11933. [PubMed: 25870862]
(13). Kim ST; Heelis PF; Okamura T; Hirata Y; Mataga N; Sancar A Biochemistry 1991, 30, 11262. [PubMed: 1958664]
(14). Vassylyev DG; Kashiwagi T; Mikami Y; Ariyoshi M; Iwai S; Ohtsuka E; Morikawa K Cell 1995, 83, 773. [PubMed: 8521494]
(15). Park HW; Kim ST; Sancar A; Deisenhofer J Science 1995, 268, 1866. [PubMed: 7604260]
(16). Vande Berg BJ; Sancar GB J. Biol. Chem 1998, 273, 20276. [PubMed: 9685377]

(17). Christine KS; MacFarlane AW IV; Yang K; Stanley RJ J. Biol. Chem 2002, 277, 38339. [PubMed: 12169694]

(18). Torizawa T; Ueda T; Kuramitsu S; Hitomi K; Todo T; Iwai S; Morikawa K; Shimada I J. Biol. Chem 2004, 279, 32950. [PubMed: 15169780]

(19). Komori H; Masui R; Kuramitsu S; Yokoyama S; Shibata T; Inoue Y; Miki K Proc. Natl. Acad. Sci. U. S. A 2001, 98, 13560. [PubMed: 11707580]

(20). Mees A; Klar T; Gnau P; Hennecke U; Eker AP; Carell T; Essen LO Science 2004, 306, 1789. [PubMed: 15576622]

(21). Kiontke S; Geisselbrecht Y; Pokorny R; Carell T; Batschauer A; Essen LO EMBO J. 2011, 30, 4437 [PubMed: 21892138]

(22). Kim ST; Heelis PF; Sancar A Biochemistry 1992, 31, 11244. [PubMed: 1445863]

(23). Prytkova TR; Beratan DN; Skourtis SS Proc. Natl. Acad. Sci. U S. A 2007, 104, 802. [PubMed: 17209014]

(24). Lee W; Kodali G; Stanley RJ.; Matsika S Chem. - Eur. J 2016, 22, 11371. [PubMed: 27362906]

(25). Antony J; Medvedev DM; Stuchebrukhov AA J. Am. Chem. Soc 2000, 122, 1057.

(26). Medvedev D; Stuchebrukhov AA J. Theor. Biol 2001, 210, 237. [PubMed: 11371177]

(27). Kao YT; Saxena C; Wang L; Sancar A; Zhong D Proc. Natl. Agd Sci. U S. A 2005,102,16128.

(28). Wang H; Chen X; Fang W Phys. Chem. Chem. Phys 2014, 16, 25432. [PubMed: 25341360]

(29). Liu ZY; Tan C; Guo XM; Kao YT; Li J; Wang LJ; Sancar A; Zhong DP Proc. Natl. Acad. Sci. U. S. A 2011, 108, 14831. [PubMed: 21804035]

(30). Liu ZY; Guo XM; Tan C; Li J; Kao YT; Wang LJ; Sancar A; Zhong DP J. Am. Chem. Soc 2012, 134, 8104–8114. [PubMed: 22533849]

(31). Narayanan M; Singh VR; Kodali G; Moravcevic K; Stanley RJ Photochem. Photobiol 2017, 93, 343. [PubMed: 27935052]

(32). Weber S; Mobius K; Richter G; Kay CWM. J. Am. Chem. Soc 2001, 123, 3790. [PubMed: 11457111]

(33). Migliore A J. Chem. Theory Comput 2011, 7, 1712. [PubMed: 26596435]

(34). Tamada T; Kitadokoro K; Higuchi Y; Inaka K; Yasui A; de Ruiter PE; Eker AP; Miki K Nat. Struct. Mol. Biol 1997, 4, 887.

(35). Fujihashi M; Numoto N; Kobayashi Y; Mizushima A; Tsujimura M; Nakamura A; Kawarabayasi Y; Miki K J. Mol. Biol 2007, 365, 903. [PubMed: 17107688]

(36). Kuan D; Duff S; Posarac D; Bi X Can. J. Chem. Eng 2015, 93, 640.

(37). Trott O; Olson AJ J. Comput. Chem 2010, 31, 455. [PubMed: 19499576]

(38). Case DA; Darden TA; Cheatham TE III; Simmerling CL ; Wang J; Duke RE; Luo R; Walker RC; Zhang W; Merz KM ; Roberts B; Hayik S; Roitberg A; Seabra G; Swails J; Gotz AW ; Kolossvary I; Wong KF; Paesani F; Vanicek J; Wolf RM; Liu J; Wu X; Brozell SR; Steinbrecher T.; Gohlke H; Cai Q; Ye X ; Wang J; Hsieh MJ.; Cui G; Roe DR; Mathews DH; Seetin MG; Salomon-Ferrer R; Sagui C; Babin V; Luchko T; Gusarov S ; Kovalenko A; Kollman PA AMBER 12; University of California: San Francisco, 2012.

(39). Sato R; Kitoh-Nishioka H; Ando K; Yamato T Chem. Phys. Lett 2015, 633, 247.

(40). Joung IS; Cheatham TE III J. Phys. Chem. B 2008, 112, 9020. [PubMed: 18593145]

(41). Jorgensen WL; Chandrasekhar J; Madura JD; Impey RW; Klein ML J. Chem. Phys 1983, 79, 926.

(42). Miyamoto S; Kollman PA J. Comput. Chem 1992, 13, 952.

(43). Essmann U; Perera L; Berkowitz ML; Darden T; Lee H; Pedersen LG J. Chem. Phys 1995, 103, 8577.

(44). Martyna GJ; Tobias DJ; Klein ML J. Chem. Phys 1994, 101, 4177.

(45). Feller SE; Zhang YH; Pastor RW; Brooks BR J. Chem. Phys 1995, 103, 4613.

(46). Phillips JC; Braun R; Wang W; Gumbart J; Tajkhorshid E; Villa E; Chipot C; Skeel RD; Kale L; Schulten K J. Comput. Chem 2005, 26,1781. [PubMed: 16222654]

(47). Beratan DN; Onuchic JN Photosynth. Res 1989, 22, 173. [PubMed: 24424807]

(48). Onuchic JN; Beratan DN J. Chem. Phys 1990, 92, 722.

(49). Betts JN; Beratan DN; Onuchic JN J. Am. Chem. Soc 1992, 114, 4043.

(50). Beratan DN; Betts JN; Onuchic JN J. Phys. Chem 1992, 96, 2852.

(51). Beratan DN; Betts JN; Onuchic JN Science 1991, 252, 1285. [PubMed: 1656523]

(52). Beratan DN; Onuchic JN; Betts JN; Bowler BE; Gray h. b. J. Am. Chem. Soc 1990,112, 7915.

(53). Balabin IA; Hu XQ; Beratan DN J. Comput. Chem 2012, 33, 906. [PubMed: 22298319]

(54). Jones ML; Kurnikov IV; Beratan DN J. Phys. Chem. A 2002, 106, 2002.

(55). Humphrey W; Dalke A; Schulten K J. Mol. Graphics 1996, 14, 33.

(56). Valiev M; Bylaska EJ; Govind N; Kowalski K; Straatsma TP; Van Dam HJJ; Wang D; Nieplocha J; Apra E; Windus TL; de Jong WA Comput. Phys. Commun 2010, 181, 1477.

(57). Frisch MJ; Trucks GW; Schlegel HB; Scuseria GE; Robb MA; Cheeseman JR; Montgomery JA Jr.; Vreven T; Kudin KN; Burant JC; Millam JM; Iyengar SS; Tomasi J; Barone V; Mennucci B; Cossi M; Scalmani G; Rega N; Petersson GA; Nakatsuji H; Hada M; Ehara M; Toyota K; Fukuda R; Hasegawa J; Ishida M; Nakajima T; Honda Y; Kitao O; Nakai H; Klene M; Li X; Knox JE; Hratchian HP; Cross JB ; Bakken V; Adamo C; Jaramillo J; Gomperts R; Stratmann RE; Yazyev O; Austin AJ; Cammi R; Pomelli C; Ochterski JW; Ayala PY; Morokuma K; Voth GA; Salvador P; Dannenberg JJ; Zakrzewski VG; Dapprich S; Daniels AD; Strain MC; Farkas O; Malick DK; Rabuck AD; Raghavachari K; Foresman JB; Ortiz JV; Cui Q; Baboul AG; Clifford S; Cioslowski J; Stefanov BB; Liu G; Liashenko A; Piskorz P; Komaromi I; Martin RL; Fox DJ; Keith T; Al-Laham MA; Peng CY; Nanayakkara A; Challacombe M; Gill PMW; Johnson B; Chen W; Wong MW; Gonzalez C; Pople JA Gaussian 03, Revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.

(58). Dunning TH J. Chem. Phys 1989, 90, 1007.

(59). Adamo C; Barone V J Chem. Phys 1999, 110, 6158.

(60). Zhao Y; Truhlar DG Theor. Chem. Acc 2008, 120, 215.

(61). Yanai T; Tew DP; Handy NC Chem. Phys. Lett 2004, 393, 51.

(62). Becke AD J. Chem. Phys 1993, 98, 5648.

(63). Casida ME; Jamorski C; Casida KC; Salahub DR J. Chem. Phys 1998,108,4439.

(64). Jensen F Introduction to Computational Chemistry, 2nd ed.; John Wiley & Sons: Chichester, 2007.

(65). Because of the non-zero overlaps, the quantities in eq 6 can be negative when they are much smaller than unity.

(66). Weissbluth M Atoms and Molecules; Academic Press: New York, 1978.

(67). Note that the value of this quantity depends on the kind of population analysis used, but such a (secondary) dependence does not affect any of the conclusions of our study.

(68). Stowasser R; Hoffmann R J. Am. Chem. Soc 1999, 121, 3414.

(69). Sancar A Biochemistry 1994, 33, 2. [PubMed: 8286340]

(70). Pauszek RF; Kodali G; Siddiqui MSU; Stanley R J. Am. Soc 2016, 138, 14880.

(71). The calculation using the M06-2X density functional was performed using the Gaussian software. Nearly identical results were obtained using the PBE0 functional in the Gaussian and NWChem packages. The NWChem software was used for the other calculations.

(72). Balabin IA; Onuchic JN Science 2000, 290, 114. [PubMed: 11021791]

(73). Troisi A; Nitzan A; Ratner MA J. Chem. Phys 2003, 119, 5782.

(74). Skourtis SS; Balabin IA; Kawatsu T; Beratan DN Proc. Natl Acad. Sci. U. S. A 2005, 102, 3552. [PubMed: 15738409]

(75). Klinman JP Chem. Phys. Lett 2009, 471, 179. [PubMed: 20354595]

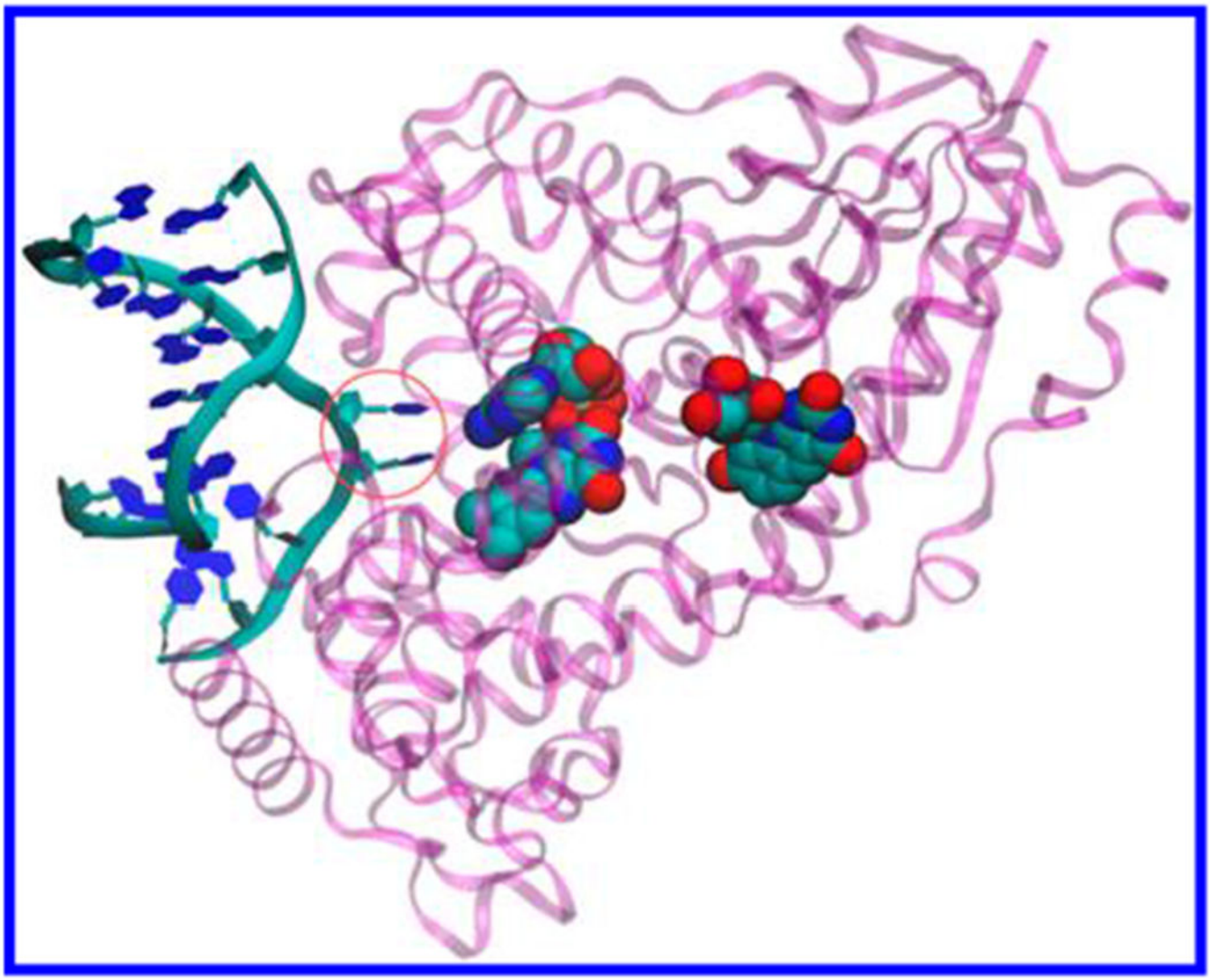


Figure 1.

DNA photolyase from *A. nidulans* (purple) complexed with repaired DNA analog (left) (crystal structure from the PDB file 1TEZ[20]). ET from the excited $FADH^-$ cofactor (CPK representation, center) splits the two thymines in CPD (circled blue hexagons). The 5-deazaflavin antenna cofactor is also shown (CPK representation, right).

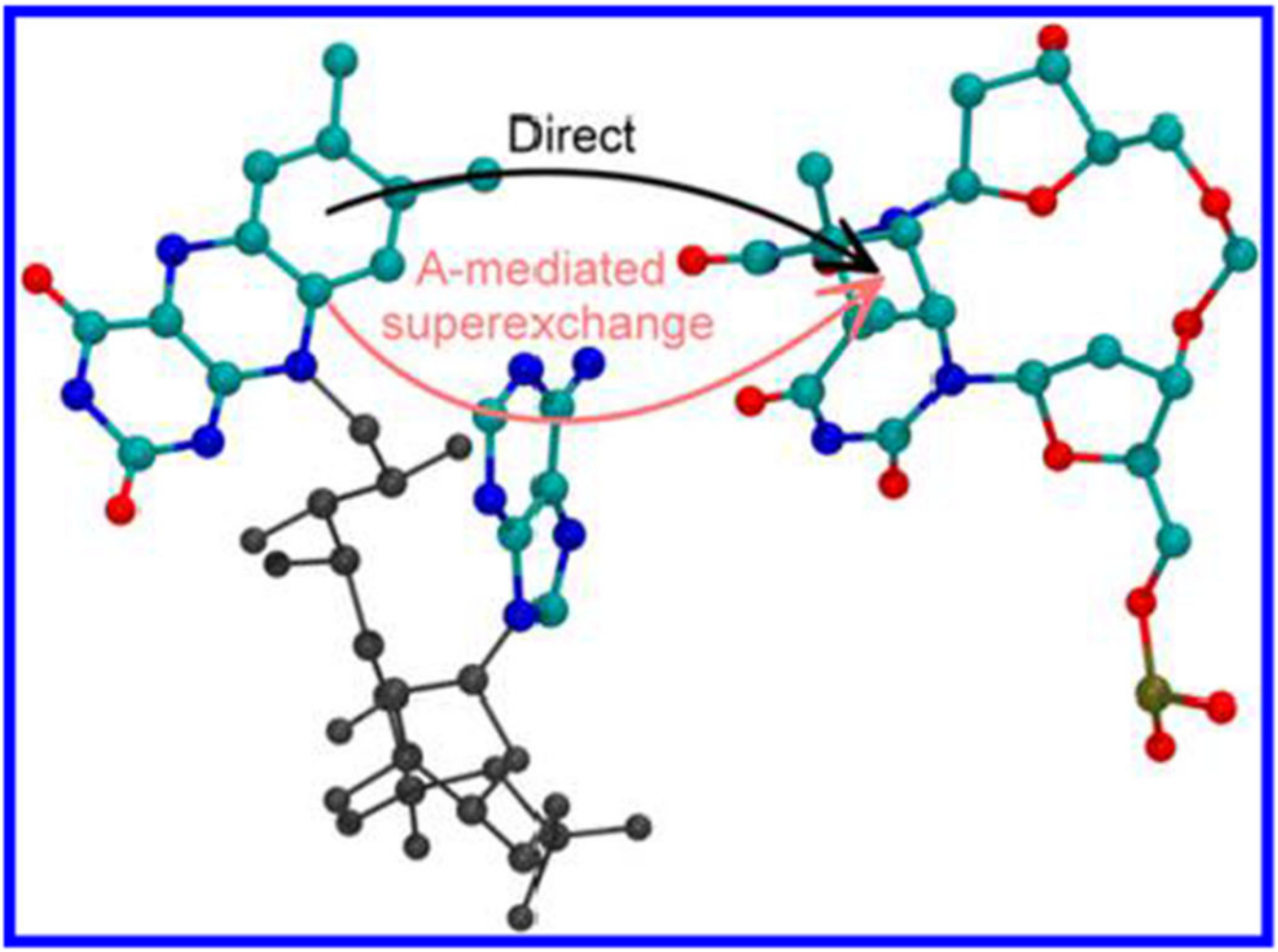


Figure 2.

Geometry of $FADH^-$ and CPD in *E. coli* photolyase (from ref 25). The arrows indicate two possible mechanisms for forward ET: direct (not adenine-mediated, in black) ET from flavin to CPD, which includes through-space tunneling and water-mediated tunneling (surrounding water molecules are not shown) and adenine-mediated superexchange (red).

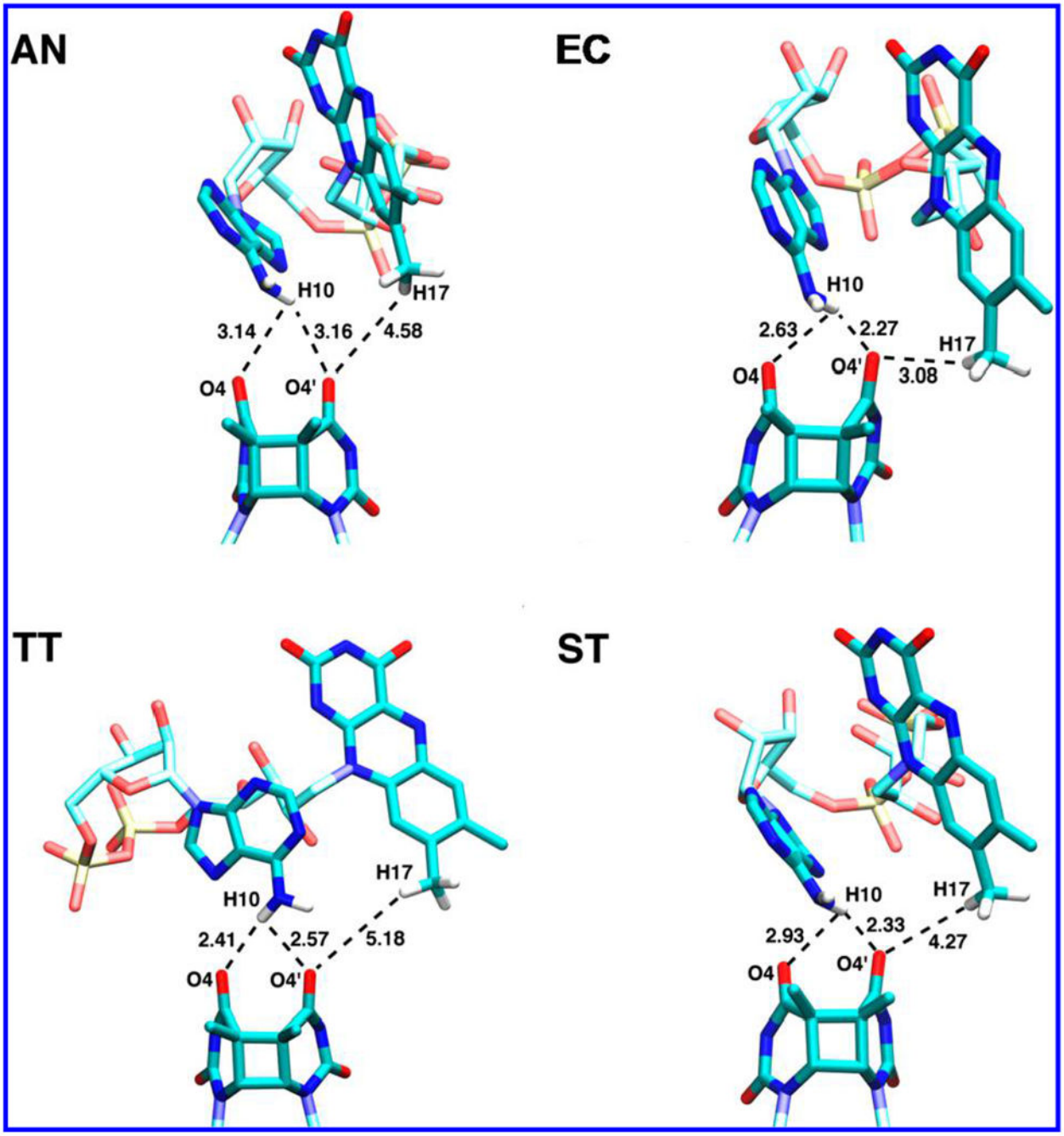


Figure 3.

Docked $FADH^{-}$–CPD conformations used in the ET study: the structures of the photolyase–CPD complex in *E. coli* (EC) from ref 25, and the crystal structures produced by our docking protocol for *A. nidulans* (AN), *T. thermophilus* (TT), and *S. tokodaii* (ST). Only the hydrogen atoms at the $FADH^{-}$–CPD interface are shown in this picture.

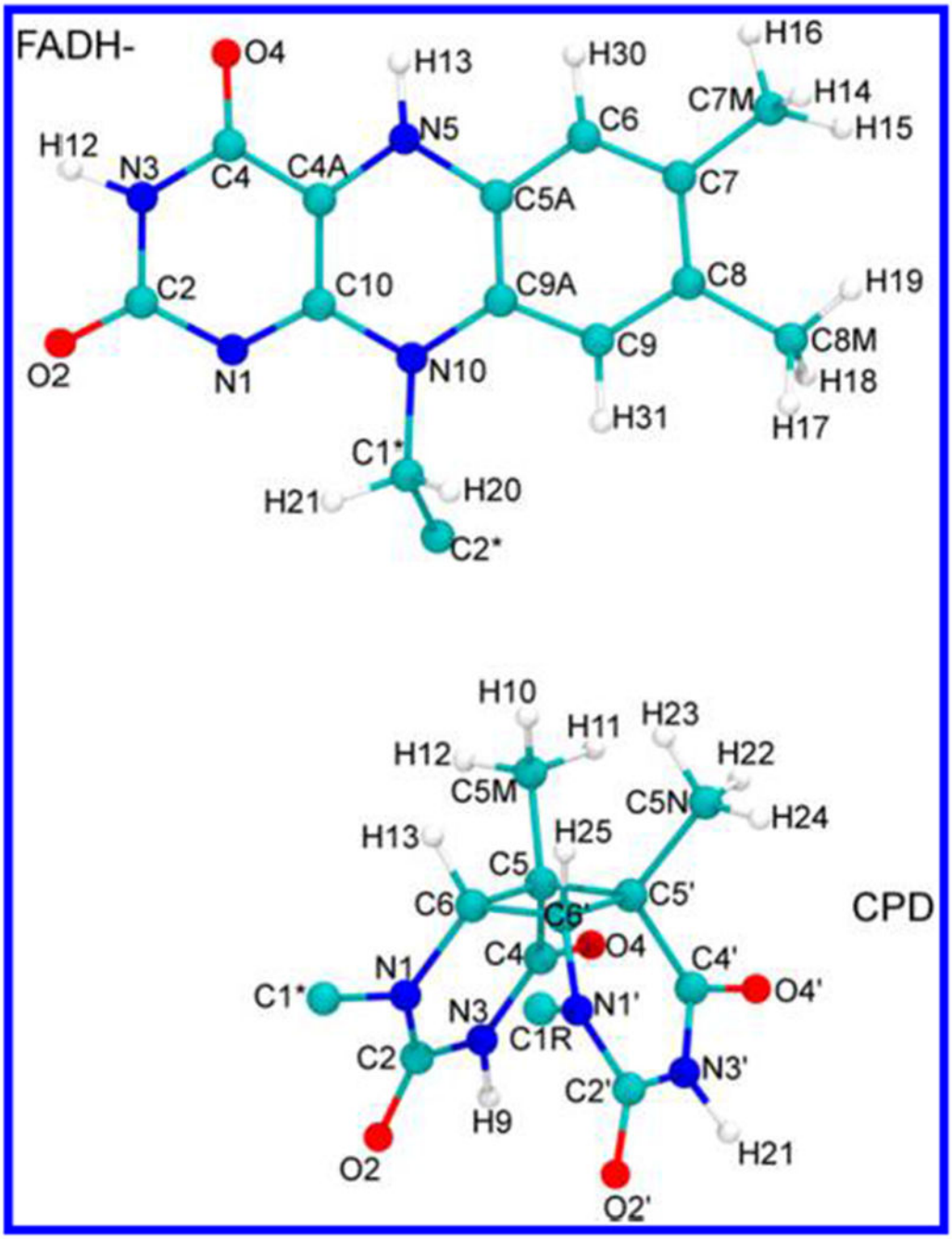


Figure 4.

Pruned flavin (above) and CPD (below) used for DFT and TDDFT calculations. The dangling C2* atom in $FADH^{-}$, and the C1*, C1R atoms in CPD, were replaced with hydrogen atoms. The molecular geometries were optimized at the B3LYP/6-31g* level before carrying out the single-point electronic structure calculations described in the text.

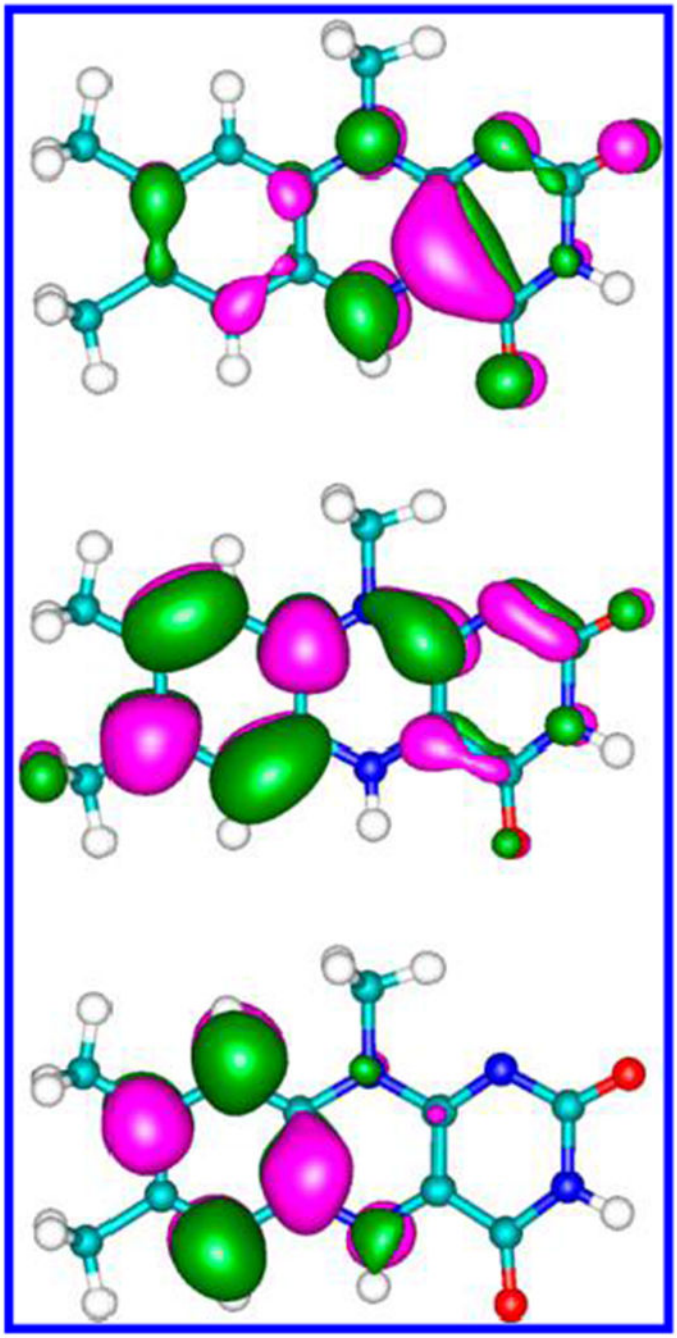


Figure 5.
HOMO (top), LUMO (middle), and LUMO+1 (bottom) of the flavin, calculated using DFT with the M06-2X density functional and the cc-pVTZ basis set.

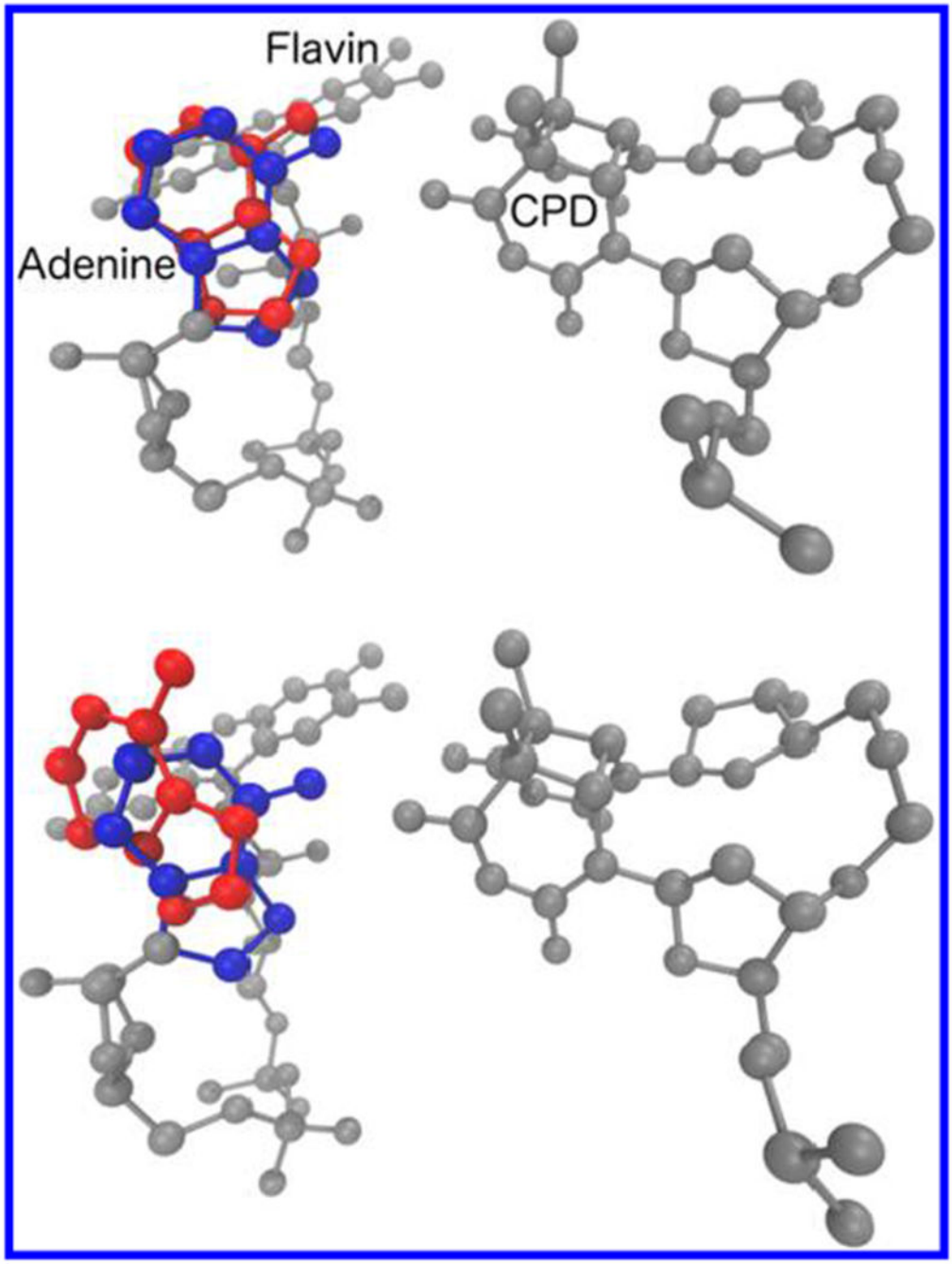


Figure 6.

Tilted conformation of *E. coli* adenine (red) at 310 K (bottom) and 293 K (top) compared to the typical proximal conformation (blue) seen in the other DNA photolyases.

Table 1.

Summary of Photolyases Studied

organism	optimal growth temperature (°C)	simulation temperature (°C)	classification
A. nidulans (PDB: 1QNF[34])	unknown[a]	20[b]	mesophile
E. coli (PDB: 1DNP[15])	37	20 and 37	mesophile
T. thermophilus (PDB: 1IQR[19])	~60	60	thermophile
S. tokodaii (PDB: 2E0I[35])	80	80	hyperthermophile

[a]The *A. nidulans* crystal structure in the 1QNF file was obtained for the PCC 6301 strain of *S. elongatus*. Recently, an optimal growth temperature of 33 °C was reported for the PCC 7942 strain of this bacterium.[36]

[b]Crystallization temperature.

Table 2.

Vertical Excitation Energy (ΔE) and Oscillator Strength (OS) of S1 and S2 Flavin States, from TDDFT Calculations with the Indicated Density Functionals[71][a]

	S1		S2	
TDDFT method	**Δ*E***	**OS**	**Δ*E***	**OS**
B3LYP + ‘00	2.79	10^{-4}	3.15	0.123
	444		394	
PBE0	2.89	2×10^{-4}	3.26	0.132
	429		380	
M06-2X	3.17	10^{-3}	3.62	0.168
	391		342	
CAM-B3LYP	3.20	2×10^{-4}	3.63	0.166
	387		342	
experimental wavelength[69,70] (nm)	~425		~360	

[a]For each method, the excitation energies are given first, in eV units, with the corresponding wavelengths, in nm, below.

Table 3.

f Values, in Percentage Form, for *A. nidulans* (AN), *E. coli* (EC), *T. thermophilus* (TT), and *S. tokodaii* (ST) Photolyases at the Listed Temperatures (in K), with the Flavin Initially in the S1 or S2 State

	species, temperature				
	AN, 293	**EC, 293**	**EC, 310**	**TT, 333**	**ST, 353**
from 25 ns of MD production run					
f(S1)	32%	28%	3%	33%	36%
f(S2)	30%	23%	2%	32%	34%
from 50 ns of MD production run					
f(S1)	33%	26%	5%	30%	31%
f(S2)	31%	21%	3%	28%	29%