Light-induced Charge Transfer in Proteins of the Cryptochrome/Photolyase Family

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

(Dr. rer. nat)

Fakultät für Chemie und Biowissenschaften

Karlsruher Institut für Technologie (KIT) - Universitätsbereich

genehmigte

DISSERTATION

von

Dipl. Phys. Gesa Lüdemann

aus

Frankfurt am Main

Dekan: Prof. Dr. P. Roesky

Referent: Prof. Dr. M. Elstner

Korreferent: Prof. (apl.) Dr. W. Wenzel

Tag der mündlichen Prüfung: 17.07.2015

Zusammenfassung

In der vorliegenden Arbeit wurden lichtinduzierte Ladungstransferprozesse in Proteinen der Familie der Cryptochrome und Photolyasen untersucht. Für deren Modellierung werden Multi-Skalen Methoden benötigt, die klassische sowie quantenchemische Ansätze verbinden. Photolyasen verwenden Elektronentransfer um DNA zu reparieren, die durch UV-Licht geschädigt wurde, und verhindern so Mutation und Absterben von Zellen. Elektronentransfer entlang einer Reihe evolutionär konservierter Tryptophane ermöglicht Cryptochromen Signale weiterzuleiten, die zum Beispiel Pflanzen helfen, sich an veränderte Lichtverhältnisse anzupassen.

Die Aktivierung von Photolyasen des Organismus *E. coli* ist experimentell gut untersucht und eignete sich damit besonders zur Weiterentwicklung der in unserer Arbeitsgruppe entwickelten Methode zur Simulation von Elektronentransfer. Der dominierende Elektronentransferpfad konnte unter Einbeziehung von Alternativpfaden eindeutig indentifiziert werden. Des weiteren wurde die Beteiligung eines vierten Tryptophans am Elektronentransfer vorhergesagt, die durch experimentelle Untersuchungen bestätigt wurde. Durch Einbeziehung elektronischer Polarisation mittels eines einfachen Skalierungsparameters, konnte die Voraussage der Transferraten verbessert werden. Der vermutete Einfluss von Nichtgleichgewichtsprozessen wurde durch die Analyse der Verweilzeiten der Tryptophan Radikale bestätigt. Die Simulationsmethode ist daher gut zur Aufklärung von experimentell noch nicht charakterisierten Cryptochromen und Photolyasen geeignet. Bei der Aktivierung von Cryptchromen des Organismus *A. thal.* ist neben dem Elektronen- auch ein Protonentransfer wichtig. Anwendung der Elektronentransfermethode sowie Berechnungen zur Freien Energie konnten zur Aufklärung der Zeitskalen der beiden Reaktionen beitragen. Die Wechselwirkung von Magnetfeldern mit Radikalpaaren, wie sie bei der Fotoaktivierung geformt werden, bildet die Basis für die Vermutung, dass Cryptochrome für die Fähigkeit von Zugvögeln, sich im Erdmagnetfeld zu orientieren, verantwortlich sind. Es konnte gezeigt werden, dass *A. thal.* Cryptochrome, wie *E. coli* Photolyase, ein langlebiges Radikalpaar bildet, das die nötigen Eigenschaften aufweist, um im Prinzip von einem Magnetfeld beeinflusst werden zu können. Eine weitere wichtige Beobachtung war ein feiner struktureller Unterschied zwischen *A. thal.* Cryptochrome und *E. coli* Photolyase, welcher als Erklärung für die experimentell beobachtete, unterschiedlich hohe Sensitivität der Proteine auf ein äußeres Magnetfeld, vorgeschlagen wurde.

Weiterhin wurden neue Modelparameter zur Simulation der Reparatur von DNA entwickelt. Es konnte gezeigt werden, dass die verwendete approximative Quantenchemie-Methode zur Beschreibung der Energien der beteiligten Moleküle geeignet ist und damit eine Methode zur Verfügung steht, mit der die Dynamik der zugrundeliegenden, atomaren Mechanismen des Reparaturprozesses, im Detail untersucht werden kann. Dadurch kann unter anderem Aufschluss über die zur Zeit diskutierte Effizienz des Reperaturmechanismuses gewonnen werden, sowie zur Auflösung der Diskrepanz zwischen experimentellen und theoretischen Ergebnissen bezüglich der Rate der Ringöffnung beitragen.

Abstract

In this work, light-induced charge transfer reactions in proteins of the photolyase/cryptochrome family are investigated using multi-scale methods which combine quantum chemical and molecular mechanics approaches.

Photolyases are enzymes which use electron transfer reactions to repair radiation damaged DNA, thereby preventing mutation and cell death. Long-range electron transfer along a chain of conserved tryptophan residues provides the basis for the signaling function of the closely related cryptochrome proteins. They mediate for example the blue-light response of plants.

The photo-activation of *E. coli* photolyase is well investigated and provided an ideal test case to further develop the electron transfer simulation approach developed in our group. We showed that the method used captures the non-equlibrium character of the reaction which is not correctly described by standard electron transfer theory. The main electron transfer pathway was identified among alternative routes and the contribution of a fourth tryptophan to the electron flow was predicted correctly. Inclusion of electronic polarization effects in an approximate fashion yielded improved electron transfer rates. Altogether, the findings proved the direct electron transfer approach to be a valuable predictive tool which can be used to establish electron transfer pathways in proteins of photolyase/cryptochrome family which are less well understood.

The photo-activation of *A. thal.* cryptochrome involves both electron and proton transfers. Direct electron transfer and potential of mean force calculations could add to solving the discrepancy regarding the rates of proton and electron transfer. The flavin-tryptophan radical pair that is formed during the photo-activation can interact with an external magnetic field. This effect is thought to provide the basis for the magnetic compass sense of migratory birds. It is shown that a persistent radical pair is formed in *A. thal.* cryptochrome, comprising all prerequisites to for its potential sensitivity to weak magnetic fields. In addition, a subtle structural difference between *E. coli* photolyase and *A. thal.* cryptochrome was identified and suggested to provide a qualitative answer to the different magnitudes of the magnetic field effect observed for the two proteins in experiment.

Furhermore, model parameters for the simulation of DNA repair by CPD photolyase were developed. It was demonstrated that the energetics of the molecules involved in the electron transfer are well described by the approximate quantum chemical method used. The developed approach is well suited to analyze the dynamic effects such as solvent polarization and environmental relaxation on an atomic level and might provide new insight in the nature of the repair reaction. Open questions, e.g. regarding the efficiency of the reaction might be investigated. In addition, future simulations might help to solve the present disaccord between the experimental and theoretical results regarding the rate of ring opening.

Contents

CHAPTER 1

Introduction

Sunlight is essential for all life on earth. It is being converted into chemical energy by plants and other photoautotrophs via the process of photosynthesis. Beyond serving as energy source, various organisms rely on sunlight to gain information about their environment and to respond to it. They do so by means of light-sensitive proteins which provide signals for functions such as the development and movement of the organism, for synchronization with circadian and seasonal rhythms or for visual perception.

On the other hand, irradiation by sunlight of a wavelenght below 300 nm is extremely harmful to DNA which is the key molecule that encodes life on earth [1]. If the photo-products resulting from UV irradiation in DNA are not repaired, they accumulate to an extend which ultimately leads to cell death [2]. To avoid this catastrophic effect of UV irradiation, all organisms possess DNA repair systems which are able to recognize and remove UV photo-products. In higher eukaryotes, as in humans, most of the UV irradiation induced lesions are typically repaired by the removal of a whole stretch of DNA around the lesions. Other organisms and in particular plants possess DNA repair enzymes which are able to repair the most common UV-induced lesions directly, splitting them back into the undamaged original nucleobases. DNA photolyases [3] are probably the most efficient and ancient ones [4].

A closely related group of enzymes has evolved from photolyases, the cryptochromes [5]. All proteins of the cryptochrome/photolyase family share considerable sequence similarity in their N-terminal photo-reactive domain, exhibit a homologous three-dimensional fold and bind the chromophore flavin adenine dinucleotide (FAD) as a cofactor [6]. Cryptochromes have mostly lost the DNA repair ability but regained a role in light-dependent signaling instead [6]. They are involved in processes such as the blue-light responses of plants [7] and have been proposed to be the long-sought receptor which allows birds and other animals to navigate in the earth' mag-

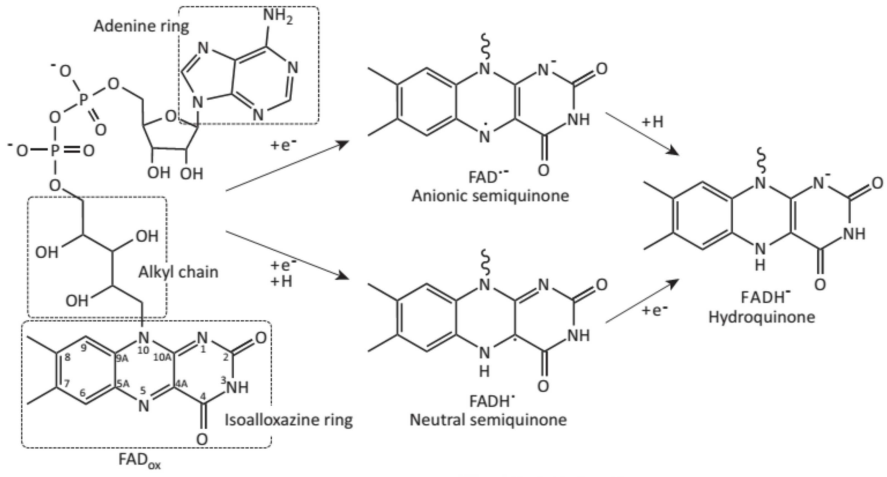


Figure 1.1: Chemical structure of FAD and of the isoalloxazine ring in its biologically relevant oxidation states. *Photolyases contain fully reduced $FADH^-$ in the active state, while cryptochromes harbour oxidized FAD_{ox}. During photo-activation and photo-repair FAD is found to cycle through different oxidation states. Reprinted from Trends in Biochemical Sciences, 38, C.A. Dodoson, P.J. Hore, M. I. Wallace, A radical sense of direction: signaling and mechanism in cryptochrome, magnetoreception, 38, 435-445, ©2013, with permission from Elsevier.*

netic field [8]. Cryptochromes also assume a light-independent role in the circadian rythm in vertebrates [9] and insects [10]. Recently, an additional class of cryptochromes was identified, the DASH-cryptochromes [11] which have been found to repair single stranded DNA [12, 13]. Photolyases occur in all organisms except of placental mammals, for example humans, while cryptochromes are mostly present in higher plants and most animals [6]. Both cryptochromes and photolyases rely on photo-induced electron transfer to efficiently carry out their function. The mechanisms developed by nature for rapid long range electron transfer within biological macromolecules could in principle serve as templates for technological applications ranging from nanoelectronics based on molecular wires to renewable energy conversion devices.

In the following, the current knowledge about the mechanism of function of photolyases and cryptochromes is summarized. The focus is put on on the molecular mechanisms associated with the electron transfer reactions which are investigated by means of computational models in this work. The theoretical fundamentals and the challenges faced upon simulation of electron transfer in biomolecules are deferred to the next chapter.

1.1 Photolyases

The UV-irradiation of DNA leads to the formation of covalent bonds between adjacent pyrimidine bases. These defects hinder DNA processing and thus lead to mutation and ultimately to cell death [2] if not repaired. Cyclobutane pyrimidine dimers (CPD) are the most common photolesions which are repaired by photolyases, being formed *in vivo* by a ratio of 3:1 [14] compared to pyrimidine-pyrimidone (6-4) photo-products. Photolyases are classified either as CPD photolyases or (6-4) photolyases, depending on the type of DNA lesion they repair. As part of this work, model parameters for the simulation of DNA repair by the former are developed. Therefore, the current knowledge of the repair mechanism of CPD photolyases is explained in the following.

1.1.1 DNA Repair by CPD Photolyase

The first crystal structures of CPD photolyase from *E. coli*[15] and *Anacystis nidulans* (*A. nidulans*) [16] revealed a large cavity which serves as the binding site for the CPD lesion. Crystallization of photolyase in complex with DNA containing a repaired CPD lesion [17] showed that the whole CPD lesion is flipped out of the DNA into the cavity. The CPD in the enzyme pocket is stabilized by the maximum number of hydrogen bonds to neighboring amino acid side chains which ensure that the CPD comes into close contact with the flavin. The structure of photolyase in complex with DNA is shown in Fig 1.2.

In all photolyases the functional form of flavin is the fully reduced form ($FADH^-$) [19]. Photolyases also contain a second cofactor which is either methenyltetrafolate (MTHF) or 8-hydroxydeazariboflavin (HDF). The second cofactor serves as antenna molecule which transfers additional excitation energy to $FADH^-$ via dipole-dipole interaction [20].

The observations of several experimental and computational groups led to the prediction of a reaction scheme for the CPD repair by photolyase which is summarized in Fig. 1.3. Direct evidence for the intermediate states and proof of the proposed photocycle came from subpicosecond transient absorption measurements[21] and femtosecond spectroscopy experiments which monitored the reaction in real time [22]. Subsequent to the excitation with blue-light, an electron is transferred from the flavin $\pi \leftrightarrow \pi*$singlet state to the pyrimidine dimer, yielding a the neutral flavin radical $FADH^\bullet$ and anionic dimer $CPD^{\bullet -}$. The C5-C5' bond in the anionic dimer breaks spontaneously followed by or concerted with the splitting of the C6-C6' bond, resulting in an intact thymine and a thymine anion radical. The electron is then returned to the transiently formed radical $FADH^\bullet$ to restore the reduced form of the enzyme, resulting in repaired DNA. The reaction rates for the individual steps of the photocycle in Fig. 1.3 are taken from the femtosecond spectroscopy study by Liu *et al* [23].

Still, some open question remain. There is unresolved disaccord between the experimental and

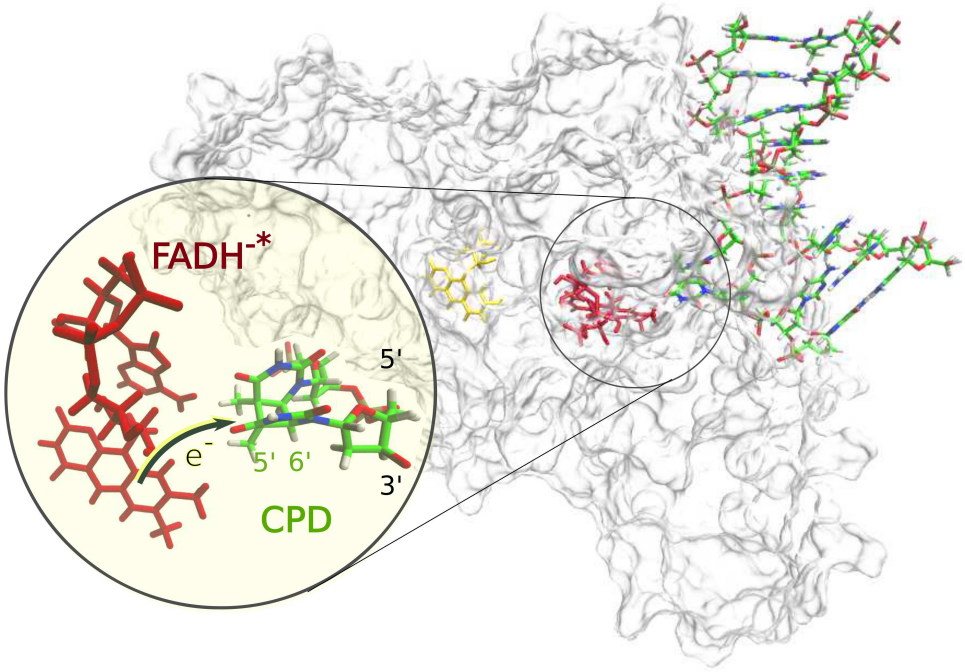


Figure 1.2: ***Structure of photolyase in complex with DNA.*** *Based on the crystal structure of A. Nidulans[16] photolyase with its flavin (red) and HDF (yellow) cofactors and bound DNA containing a CPD lesion. The thymine dimer is flipped out and inserted into the active site. As shown in the close up view, upon light excitation an electron is transferred from the flavin to CPD lesion which ultimately results in repaired DNA. The figure was prepared using VMD [18].*

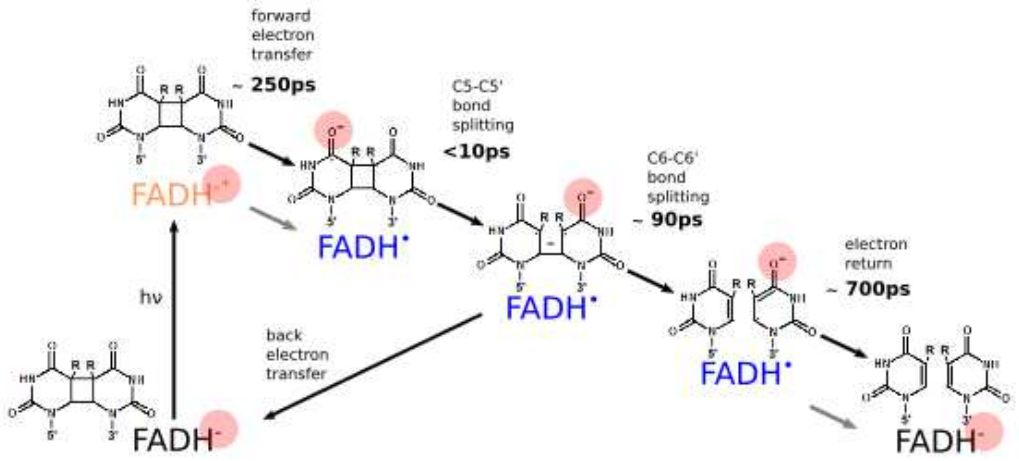


Figure 1.3: Photocycle of CPD repair by DNA Photolyase. *The lifetimes of the intermediates are taken from a femtosecond spectroscopy study Liu et al [23]. The time constant for unproductive back electron transfer is currently debated.*

computational studies regarding the rate of ring opening. While experimental data modeled the complete ring opening to occur in 100-300 ps [23, 24], computational studies predicted an spontaneous, ultrafast bond cleavage [25–27]. Also, the rate of unproductive electron back transfer and the exact quantum yield of the photo-repair reaction are currently debated [28, 29].

The electron transfer from $FADH^{-*}$ to CPD has been studied by several experimental and theoretical groups yielding inconsistent results regarding the exact mechanism of transfer. The first simulations, based on a molecular dynamics study, found the thymine dimer to be strongly coupled to the adenine moiety of the $FADH^-$ and suggested that the electron is transferred via superexchange tunneling from the flavin to the thymine dimer mediated by the adenine moiety of the FAD [30, 31]. Another theoretical model provided the conformationally averaged electronic coupling between the excited state of $FADH^-$ and the thymine dimer which indicated that the photoexcitation moves the electron closer towards the thymine dimer [32]. This localization was found to enhance the electronic coupling between flavin and CPD, suggesting that a direct through-space electron transfer from the $FADH^-$ to the CPD is the most favorable pathway. Explicit simulation of the electron dynamics based on a quantum mechanical active site model implied that the adenine moiety of FAD promotes the electron flow by providing the correct electrostatic environment [33]. Femtosecond time resolved fluorescence experiments excluded a direct role for adenine but suggest that the electron passes adenine via quantum tunneling[34], in accord with the earliest theoretical study. The present discussion is further fueled by a recent QM/MM study which indicates that the electron transfer is accompanied by proton transfer, suggesting yet another transfer mechanism [35].

1.1.2 Photo-activation of Photolyases

After purification of DNA photolyase in *in vitro* experiments the flavin cofactor exists in the inactive oxidized or semi-reduced state. The anionic, fully reduced state which is required for the repair function, can be re-established by electron transfer along a chain of conserved tryptophan residues from the protein surface to the buried cofactor. This process is induced by excitation of the flavin using red light and is referred to as photo-reduction or photo-activation. The physiological significance of this process has been questioned[36] but conservation of the three tryptophan residues in all proteins of the photolyase/cryptochrome family points to a functional role in vivo [37]. The first experiments found that in *E. coli* photolyase the excited $FADH^{\bullet}$ is reduced by intra-protein electron transfer from W306 in about 1 μs [38]. The time constant was ascribed to electron transfer by super-exchange coupling via the intermediate W382 and W359 by a theoretical study [39]. Time resolved absorption measurements showed that the $FADH^{\bullet}$ in *E. coli* is reduced much faster, in about 30 ps and attributed the previously found slow reaction to the deprotonation of the W306 radical cation [40]. Based on these findings and supported by mutational studies [41–43], stepwise hopping along the tryptophan chain, referred to as tryptophan triad, was suggested as the electron transfer mechanism. Results of polarized transient absorption spectroscopy allowed to deduct rates for the all the steps of the photo-activation reaction in semiquinone photolyase, finding the electron transfer along the triad to be completed within 30 ps, followed by deprotonation of W306 within 200 ns [44, 45]. The direct evidence for the transient oxidation of W382 and W359 came from femtosecond spectroscopy experiments which also showed the electron transfer dynamics and efficiency to be dependent on the flavin oxidation state [46, 47]. In anionic semiquinone photolyase a fourth tryptophan residue was found to be involved in addition to the dominant photo-reduction route along the tryptophan triad.

1.2 Cryptochromes

Many of the numerous physiological functions of cryptochromes are known as well as their signaling partners in the cell [6, 7]. It has for example been shown that signaling molecules which mediate photomorphogenesis and flowering [48, 49] bind to light activated *Arabidopsis thaliana* (*A. thal.)* cryptochromes. The biological activation, on the other hand, is less well understood. The proposed mechanism is that the light-driven photo-reduction leads to a conformational change of the C-terminal domain which makes cryptochrome accessible to the signaling partners. This is supported for example by observations of a conformational change related to a redox-reaction for cryptochromes from *A. thal.* and *Drosophila melanogaster* (*D. mel.*) [50, 51]. Return to the inactive state is suggested to be achieved by re-oxidation involving molecular oxygen [52]. Light-driven cycling between alternate redox states has thus been proposed to provide the basis for

signaling in plant [53] and animal cryptochromes [54, 55].

Experiments involving mutations of the tryptophan triad resulted in decreased flavin reduction in vitro but did not seem to eliminate biological activity of cryptochrome in vivo [56, 57]. This put the physiological significance of the photo-activation reaction into question. The discrepancy is suggested to be explained by the in vivo interaction of cryptochrome with small molecule activators such as ATP [58–60] which are supposed to facilitate the use of alternate electron transfer pathways involving tyrosine residues [60].

1.2.1 Photocycle of Plant Cryptochromes

In plant cryptochromes, where the photo-cycle is understood best, the neutral semiquinone $FADH^{\bullet}$ is formed upon illumination with blue light and represents the signaling state [61, 62]. In addition to the electron transfer along the tryptophan triad, protonation of the flavin and deprotonation of the tryptophan radical has been observed [63–65]. Fourir Transform Infra-Red (FTIR) spectroscopy suggests that the proton donor is an aspartic acid located in the close vicinity of the flavin [63]. In a femtosecond spectroscopy study the electron transfer in a plant cryptochrome from the green alga *Chlamydomonas reinhardtii* is found to occur in less than 100 ps while flavin protonation was detected in the μs time domain [64]. Deprotonation of the tryptophan radical was observed within picoseconds, most likely to a histidine resdiue [66].

Most recently, Müller and coworkers found that the rate of flavin protonation in *A. thal.* cryptochrome can be modulated by ATP binding and pH [58, 59]. In the case of high pH and ATP absence they observed some ultrafast protonation of flavin, although the probability for millisecond flavin protonation was still found to be higher. A QM analysis of the active site in *A. thal.* cryptochrome [67] had also suggested flavin protonation to happen ultrafast. Transient absorption spectroscopy observed deprotonation of the tryptophan radical within microseconds for *A. thal.* cryptochrome and *E. coli* photolayse *in vitro* in water-glycerol mixture [68].

1.2.2 Cryptochromes and Magnetoreception

A fascinating, yet speculative hypothesis is that cryptochromes constitute the long-sought receptor by which migratory birds perceive the Earth's magnetic field [8]. Evidence for the hypothesis has been collected by researchers from different disciplines, with zoologists and neurophysiologists on the one hand and spin chemist on the other but conclusive validation is lacking. The results of their studies are reviewed in detail by [69–72].

The proposed sensitivity of cryptochromes to the Earth's magnetic fields is based on the radical pair mechanism [69, 73, 74]. Originally suggested by Schulten [75], the magnetic sensitivity of radical pair reactions in *in vitro* was demonstrated [73]. The spin-correlated radical pair, e.g. the flavin-tryptophan radical pair created during the photo-activation, possess a magnetic moment

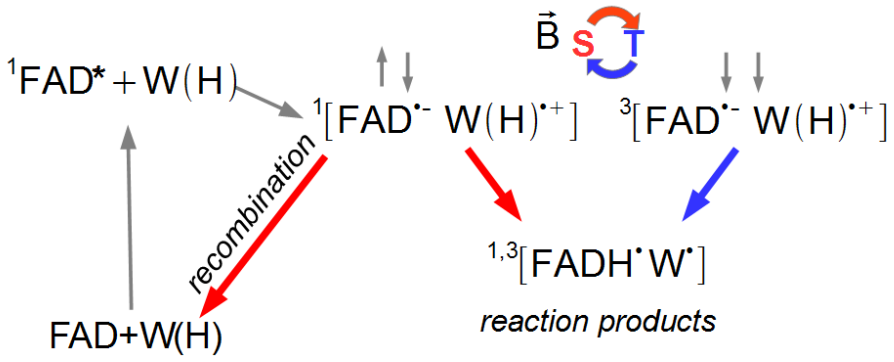


Figure 1.4: Radical pair reaction scheme. *The $^1[F^\bullet\ W^\bullet]$ state is formed trough electron transfer. Arrows indicate spin-dependent radical pair reactions that can occur either from the singlet(red) or triplet(blue) state. An external magnetic field influences the singlet to triplet interconversion, thereby modulating the ratio between recombination and product formation.*

that interacts with the magnetic nuclei and and with external magnetic fields. It is thought that the interaction with the external magnetic field alters the interconversion rate between singlet and triplet states and thereby influences the lifetime of the radical pair, leading to a change of the rate constants or the product yield of the molecular species. This effect is called magnetic field effect [72].

The first direct measurements of magnetic fields resulting in an effect on the chemical rate constants of the radicals species in *A. thal.* cryptochrome and *E. coli* photolyase have been made recently by Maeda *et.al.* They found that the recombination reaction of the anionic flavin FAD^- and tryptophan radical $W^{\bullet+}$ increases in the presence of an outer magnetic field while the protonation and deprotonation reactions were suppressed by 10-20%. The magnetic field effect for *A. thal.* cryptochrome was found to be higher than for *E. coli* photolyase. The proposed reaction scheme is shown in Fig. 1.4.

1.3 Overview Over this Work

In the next chapter the theoretical fundamentals needed to study the electron transfer reactions in proteins of the cryptochrome/photolyase family are introduced. The two approaches to model biological systems, quantum chemical methods and molecular mechanics methods are presented. The most important concepts for the theoretical description of electron transfer in biological systems are introduced and the limits of state-of-the-art Marcus' theory are discussed. Subsequently, the direct electron transfer approach which is used for the simulation of electron transfer in this work is presented.

In chapter 3 the direct electron transfer approach is applied to study possible alternative pathways for the photo-activation reaction in *E. coli* photolyase. Further, open questions regarding the mechanism of electron transfer are clarified. Inclusion of the effect of electronic polarization in the direct electron transfer approach via electrostatic potential scaling is tested and its effect on the electron transfer rates is discussed.

Following, the photo-activation reaction of *A. thal.* cryptochrome is investigated using the direct electron transfer approach and potential of mean force calculations. Furthermore, the results of the radical pair formation are discussed in terms of magnetoreception in chapter 4.

In chapter 5 electron transfer parameters for aromatic side chains are re-evaluated which might be used in the search for alternate electron transfer pathways in cryptochromes.

Chapter 6 presents the development of a model for the simulation of DNA repair by CPD-photolyase. Previous approaches and their drawbacks are discussed. Further, the model parameters are tested against higher-level methods.

Finally, a conclusion in chapter 7 summarizes the findings of this work.

CHAPTER 2

Theoretical Framework and Methods

Charge transfer reactions in biomolecules are determined by effects which occur on different time and lenght scales. On the one hand the electronic structure needs to be treated explicitly via quantum chemical methods which is only possible for small systems up to 100 atoms. On the other hand charge transfer is affected by the electrostatic and dynamic effects of the complete biomolecule which might consist of as much as several 100000 atoms. These effects are captured by empirical force field models. This chapter briefly recapitulates the methods used in this work. For a more detailed account the reader is referred to the corresponding textbooks written on this matter [76–78].

Both types of methods are required for the simulation of charge transfer reactions in biomolecules. The most widely applied theory for the description of electron transfer is Marcus' theory [79, 80]. The most important concepts as well as its limitations will be explained in the following and an alternative, more direct electron transfer approach which has been developed in our group over the last years[81, 82] is presented.

2.1 Quantum Chemistry

The starting point for electronic structure calculations is usually the stationary Schrödinger Equation (SE)

$$\hat{H}\psi = E\psi \tag{2.1}$$

with $\hat{H}$ being the time-independent Hamiltonian for a molecular system of M nuclei and N electrons. Usually the Born-Oppenheimer approximation (BOA) is applied which assumes that the

motion of nuclei and electrons can be separated. Within this approximation, the Hamiltonian depends on the electronic coordinates **r** and only parametrically on coordinates of nuclei **R**:

$$\hat{H} = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{en} + \hat{V}_{nn} \tag{2.2}$$

$$= -\frac{1}{2}\sum_{i}^{N}\nabla_i^2 + \sum_{i,j>i}^{N}\frac{1}{|\mathbf{r_i}-\mathbf{r_j}|} - \sum_{i}^{N}\sum_{A}^{M}\frac{Z_A}{|\mathbf{R_A}-\mathbf{r_i}|} + \sum_{A,B>A}^{M}\frac{Z_A Z_B}{|\mathbf{R_A}-\mathbf{R_B}|} \tag{2.3}$$

where $\hat{T}_e$ denotes the kinetic energy of the electrons, $\hat{V}_{ee}$ the electron-electron interaction, $\hat{V}_{en}$ the electron-nucleus interaction and $\hat{V}_{nn}$ the nucleus-nucleus repulsion. An analytical solution to the SE is only possible for single electron systems as e.g. H_2^+. All quantum chemical methods aim to find the best possible approximation of the wavefunction ψ in order to obtain the total energy E of the molecular system. This approach is based on the variational principle which states that any trial wavefunction yields an energy expectation value which is higher than the true ground state density E_0.

2.1.1 Hartree-Fock

The Hartree approximation is based on the assumption of a system of N non-interacting electrons. The total wavefunction of the system ψ can be written as a product of single particle wave functions ϕ_i:

$$\psi(\mathbf{r_1},...,\mathbf{r_N}) = \phi_1(\mathbf{r_1})\phi_2(\mathbf{r_2})...\phi_N(\mathbf{r_N}) = \prod_{i}^{N}\phi_i(\mathbf{r_i}) \tag{2.4}$$

Assuming that each electron i is affected by the average potential created by the remaining N-1 electrons, the N electron problem reduces to N single electron problems:

$$\hat{h}_i(\mathbf{r_i})\phi_i(\mathbf{r_i}) = \epsilon_i\phi_i(\mathbf{r_i}) \tag{2.5}$$

with $\hat{h}_i(\mathbf{r_i})$ being an effective single particle Hamiltonian which acts on the single particle wavefunction $\phi_i(\mathbf{r_i})$. This so called Hartree-Approximation does not explicitly take into account the interaction between electrons, they are uncorrelated.

According to the Pauli Principle, two electrons with identical spin can not populate the same orbital. This can be accounted for by using spin orbitals

$$\chi_i(\mathbf{x}) = \begin{cases} \phi_i(\mathbf{r})\alpha(\omega) \\ \text{or} \\ \phi_i(\mathbf{r})\beta(\omega) \end{cases} \tag{2.6}$$

and by constructing an antisymmetric wavefunction which changes sign upon exchange of the

coordinates **r** of two electrons. The resulting wavefunction is usually written as a determinant, the so called Slater-Determinant ([83])

$$\psi_N = |\chi_1(\mathbf{x_1})\chi_2(\mathbf{x_2})...\chi_N(\mathbf{x_N})| \tag{2.7}$$

Applying the variational principle in order to find the set of spin orbitals χ_N which minimizes the expectation value of the Hamiltonian yields the Hartree-Fock (HF) equations:

$$\hat{f}_i\chi(\mathbf{x_i}) = \epsilon\chi(\mathbf{x_i}) \tag{2.8}$$

including the Fock operator $\hat{f}_i$ and the Hartree-Fock potential v_{HF} which are definded as

$$\hat{f}_i = \hat{h}_i + \sum_j^n (\hat{J}_j - \hat{K}_j) \tag{2.9}$$

$$= \hat{h}_i + v_{HF}. \tag{2.10}$$

$\hat{h}_i$ is the effective one-electron operator, describing the motion of one electron in the field of all nuclei

$$\hat{h}_i = -\frac{1}{2}\nabla_i^2 - \sum_A^M \frac{Z_A}{|\mathbf{R_A} - \mathbf{r_i}|}. \tag{2.11}$$

v_{HF} is the average potential experienced by electron i due to the presence of the other electrons and represents the electron-electron repulsion in an average fashion. $\hat{J}_j$ is the Coulomb operator, describing the classical coulomb repulsion between two charge densities.

$$\hat{J}_j = \langle\chi_j(\mathbf{x_i})|\frac{1}{r_{ij}}|\chi_j(\mathbf{x_i})\rangle \tag{2.12}$$

$\hat{K}_j$ is the Exchange operator and leads to an exchange of the variables in two spin orbitals and occurs due to the antisymmetry of the Slater determinant.

$$\hat{K}_j(\mathbf{x_i})|\chi_i(\mathbf{x_i})\rangle = \langle\chi_j(\mathbf{x_j})|\frac{1}{r_{ij}}|\chi_i(\mathbf{x_j})\chi_j(\mathbf{x_i})\rangle \tag{2.13}$$

Exchange contributions exist only for electrons with the same spin. The Fock operator depends on the spin orbitals χ_i but they are also its solutions. Therefore the HF problem must be solved iteratively until the field generated by the electrons is self-consistent (SCF).

From the Hartree-Fock total energy

$$E = \sum_i h_i + \frac{1}{2}\sum_{ij}\left(J_{ij} - K_{ij}\right) + V_{nn} \tag{2.14}$$

it becomes apparent that for $i = j$ the Coulomb interaction J_{ii} of the charge distribution of a single electron with itself is calculated. This so called self-interaction is unphysical. However, in HF for $i = j$ the coulomb and exchange interaction become identical and cancel each other. In DFT the self-interaction constitutes a major obstacle.

The fock eigenvalues ϵ_i represent the orbital energies of the system. Within the Frozen Orbital Approximation (the orbitals of the N electron state are assumed to be identical to the orbitals of the N±1 electron state), the fock eigenvalues provide a way to calculate approximate ionization potentials (IP) or electron affinities (EA). The difference of the total energy of a system of N electrons and the same system where an electron from orbital k is removed, equals the orbital energy ϵ_k

$$-\epsilon_k \approx E_{N-1}^k - E_N = IP. \tag{2.15}$$

The difference of the total energy of a system of N electrons and the same system where an electron is added to the orbital l is removed, equals the orbital energy ϵ_l

$$-\epsilon_l \approx E_N - E_{N+1}^l = EA. \tag{2.16}$$

This relation is known as Koopmans' Theorem[84].

2.1.2 Density Functional Theory

Density functional theory is based on the idea that the complicated N-electron problem can be reduced by replacing the N-dimensional wavefunction by the 3-dimensional electron density $\rho = \rho(r)$. Hohenberg and Kohn could indeed show that the electron density unambiguously defines the properties of the molecular system and minimizes the energy functional for the true ground state density [85]. The contributions to the DFT energy functional

$$E\left[\rho\right] = T\left[\rho\right] + V_{en}\left[\rho\right] + J\left[\rho\right] + E_{XC}\left[\rho\right] + V_{nn} \tag{2.17}$$

include the interactions of the electrons with the static external potential of nuclei $V_{en}\left[\rho\right]$ and the classical coulomb interaction between electrons $J\left[\rho\right]$ which is defined in analogy to the Hartree theory as

$$J\left[\rho\right] = \frac{1}{2}\int\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{\mathbf{r}-\mathbf{r}'} d\mathbf{r} d\mathbf{r}' \tag{2.18}$$

$$V_{en}\left[\rho\right] = \sum_A \int \frac{Z_A \rho(\mathbf{r})}{\mathbf{R_A}-\mathbf{r}} d\mathbf{r}. \tag{2.19}$$

In the formulation of the Kohn Sham (KS) Approximation [86] orbitals are reintroduced which lead to the effective one-electron equations

$$-\frac{1}{2}\nabla^2 + v_{KS}[\rho]\,\chi_i = \epsilon_i \chi_i \tag{2.20}$$

with

$$v_{KS} = \frac{\delta V_{en}[\rho]}{\delta\rho} + \frac{\delta J[\rho]}{\delta\rho} + \frac{\delta E_{XC}[\rho]}{\delta\rho} \tag{2.21}$$

being the effective KS potential. Starting from an approximate density, the Kohn-Sham equations are solved self-consistently. The resulting KS-orbitals χ_i are used to calculate the kinetic energy

$$T[\rho] = \sum_i \langle\chi_i| -\frac{1}{2}\nabla^2 |\chi_i\rangle. \tag{2.22}$$

All the contributions arising from the many body effects are covered by the exchange-correlation functional $E_{XC}[\rho]$. The Kohn-Sham approach would yield the exact energy, if the exact form of $E_{XC}[\rho]$ was known. Also, the exact exchange-correlation functional would cancel the unphysical self-interaction arising from the coulomb term. In all real applications, however, an approximation for the exchange-correlation functional is required. The search for more accurate exchange-correlation potentials is the central goal of modern quantum chemistry [76].

Several approaches to approximate $E_{XC}[\rho]$ exist. The local density approximation (LDA), where E_{XC} depends on the electron density only, is the simplest approach. Another class of functionals called general gradient approximation (GGA) depends on the electron density and its gradient. A popular representative of this class is the PBE [87] functional. However, both approaches do not describe the correct asymptotic $1/r$ behaviour of the exchange potential which is needed to cancel the unphysical self-interaction. In HF, however, the asymptotic behaviour is described correctly and the most successful exchange-correlation functionals include an exchange term from HF. These functionals are called hybrid functionals. Here, B3LYP[88] is the most widely used functional which contains about 20% contribution of HF exchange. Yet another class are the so called range-separated functionals[89] which use the exchange from the LDA/GGA approach for small distances r and the exchange from HF for larger distances.

According to Janak's theorem[90] the KS HOMO energy is directly related to the vertical ionization potential

$$-\epsilon_{HOMO} = IP_v. \tag{2.23}$$

This relation is correct only for the exact exchange-correlation potential. In practice, the KS energies differ from the experimental ionization energies depending on the exchange-correlation functional used. For LDA and GGA functionals a deviation from the experimental ionization energies by more than 2eV is found [91]. For hybrid functionals the error is reduced while range-

separated functionals [89] are a more promising solution to this problem.

The reason why the relation Eq. 2.23 is not fullfilled by any approximate exchange correlation potentials is a phenomenon called derivative discontinuity. Perdew *et al.* [92] extended the Kohn-Sham theory to fractional electron numbers and showed that the exchange-correlation potential jumps as the number of electrons passes through an integer. All current exchange-correlation potentials are characterized by a continuous potential with respect to the electron number.

2.1.3 Density Functional Tight Binding

Density functional tight binding (DFTB) is an approximative quantum chemical method which is derived from DFT. Though less accurate than DFT or *ab initio* methods on average, its main advantage is the increased computational speed by 2-3 orders of magnitude. This allows to treat large molecules and sufficiently long sampling for QM/MM molecular dynamics simulations [93].

Tight binding theories (TB) originate from solid state physics and assume that electrons are tightly bound to the nuclei. Therefore, the electron density can be expressed in terms of a a reference density ρ_0 and fluctuations $\delta\rho_0$

$$\rho(\mathbf{r}) = \rho_0(\mathbf{r}) + \delta\rho_0(\mathbf{r}). \tag{2.24}$$

In this spirit, the DFTB total energy is derived by expanding the DFT total energy Eq.(2.17) around a reference density ρ_0 up to the third order:

$$E\left(\rho_0(\mathbf{r}) + \delta\rho_0(\mathbf{r})\right) = \sum_i^{occ} \langle\psi_i| -\frac{1}{2}\nabla^2 + v_{ext} + \int \frac{\rho_0(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + v_{XC}\left[\rho_0\right] |\psi_i\rangle \tag{2.25}$$

$$+ E_{XC}[\rho_0] + E_{nn} - \int v_{XC}[\rho_0](\mathbf{r})\rho_0(\mathbf{r}) - \frac{1}{2}\int\int \frac{\rho_0(\mathbf{r}')\rho_0(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \tag{2.26}$$

$$+ \frac{1}{2}\int\int \left(\frac{\delta^2 E_{XC}[\rho]}{\delta\rho\delta\rho'} + \frac{1}{\mathbf{r}-\mathbf{r}'}\right)\delta\rho\delta\rho' d\mathbf{r}d\mathbf{r}' \tag{2.27}$$

$$+ \frac{1}{6}\int\int\int \frac{\delta^3 E_{XC}[\rho]}{\delta\rho\delta\rho'\delta\rho''} d\rho d\rho' d\rho'' d\mathbf{r}d\mathbf{r}'d\mathbf{r}'' \tag{2.28}$$

$$\tag{2.29}$$

The total DFTB energy is an approximation to Eq. 2.29 and can be written as:

$$E^{DFTB} = \sum_i \langle\psi_i| H\left[\rho_0\right] |\psi\rangle + E^{rep} + E^{2nd} + E^{3rd}. \tag{2.30}$$

Truncation of the Taylor series after the first order gives the standard non-selfconsistent DFTB method [94] which works well for solid state systems. Including second-order terms accounts for charge fluctuations which need to be considered for the description of molecules and results in

the DFTB2 method [95]. Inclusion of the third-order allows for an improved description of highly charged molecules and yields the DFTB3 method [93, 96].

The first energy term is the energy contribution from an atomic orbital Hamiltonian depending only on the reference density. The repulsive energy term contains all the core electron effects, so that only valence electrons need to be considered. As the TB theory assumes tightly bound electrons, the Kohn-Sham orbitals are expanded in a minimal basis μ:

$$\sum_i \langle \psi_i | H\left[\rho_0\right] | \psi_i \rangle = \sum_i^{occ} \sum_{\mu,\nu} c_\mu^i c_\nu^i H_{\mu\nu}^0. \tag{2.31}$$

The matrix elements of the DFTB Hamiltonian are the principal parameters of the method[97]. The diagonal elements are Kohn-Sham eigenvalues obtained from an atomic DFT calculation using the PBE[87] exchange-correlation functional and confined Slater-type basis functions. The values for the Hamilton and overlap matrix elements are precomputed and tabulated for each orbital pair and atom for different inter-atomic distances and need not to be calculated during the runtime.

The repulsive energy term is approximated in the spirit of TB theory as a sum of pairwise potentials,

$$E^{rep} = \frac{1}{2} \sum_{ab} V_{rep}^{ab} \left[\rho_0^a, \rho_0^b, r_{ab}\right] \tag{2.32}$$

where r_{ab} is the distance between atoms a and b. Three center contributions are neglected. The potentials are fitted to reproduce DFT energies, geometries and vibrational frequencies.

The second order term accounts for charge density fluctuations $\delta\rho$ which are written as superposition of atomic contributions $\delta\rho = \sum_a \delta\rho_a$. Applying the monopole approximation,

$$\delta\rho \approx \triangle q_a F_{00}^\alpha Y_{00} \tag{2.33}$$

the radial dependence of the density fluctuation on atom a is assumed to be spherical. The second-order energy can thus be written as

$$E^{2nd} \approx E^\gamma = \frac{1}{2} \sum_{ab} \triangle q_a \triangle q_b \gamma_{ab}, \tag{2.34}$$

with γ_{ab} reducing to $1/r_{ab}$ for large inter atomic distances describing a pure Coulomb interaction between the partial charge $\triangle q_a$ and $\triangle q_b$. For vanishing inter atomic distances $a = b$, γ represents the electron-electron interaction on a single atom described by the Hubbard Parameter U_a. The Hubbard parameter is twice the chemical hardness and is a measure for the energy change of a system upon addition or removal of an electron.

In the second order approximation the Hubbard parameter is independent of the charge state.

The third order term allows the Hubbard parameter, which represents the size of an atom, to be charge dependent: as charge accumulates around an atom its effective size increases and vice versa.

$$E^{3rd} = \frac{1}{3}\sum_{ab}\left(\triangle q_a\right)^2 \triangle q_b \Gamma_{ab} \tag{2.35}$$

Γ contains the derivative of the γ function with respect to the charge.

$$\Gamma_{ab} = \frac{\partial \gamma_{ab}}{q_a}|_{q_a^0} = \frac{\partial \gamma_{ab}}{U_a}\frac{\partial U_a}{q_a}|_{q_a^0} \tag{2.36}$$

$$\Gamma_{aa} = \frac{\partial \gamma_{aa}}{q_a}|_{q_a^0} = \frac{1}{2}\frac{\partial \gamma_{aa}}{U_a}\frac{\partial U_a}{q_a}|_{q_a^0} \tag{2.37}$$

The Hubbard derivative is pre-computed.

2.2 Empirical Force field Models

For large molecular systems it is not feasable to use the quantum chemical Hamiltonian as a description. Instead, empirical force field methods are used which calculate the energy of a molecular system as a function of the nuclear positions only. Based on the BOA, the electrons are assumed to instantly follow the motion of the nuclei. The nuclei are treated as spheres carrying atom-centered partial charges and their interaction is described by classical mechanics models. The empirical potential energy function which accounts for the interactions between all atoms is called "forcefield". The representative statistical ensemble from which macroscopic properties of the molecular system can be derived is generated by numerical integration of the Newtonian equations of motion.

2.2.1 Force Field Description

In this work the AMBER (Assisted Model Building and Energy Refinement) force field for proteins and nucleic acids [98] was used, which is of the functional form:

$$V_{AMBER} = V_{bonded} + V_{non-bonded} \tag{2.38}$$

$$V_{bonded} = \sum_{bonds} K_r(r-r_0)^2 + \sum_{angles} K_\theta(\theta-\theta_0)^2 + \sum_{dihedral}\frac{V_n}{2}(1+cos(n\phi-\phi_0)) \tag{2.39}$$

For the bond distance r and bond angle θ terms a harmonic potential is used to describe how the energy changes with the deviation from their respective equilibrium value. For the dihedral angle ϕ term a torsional potential is used to model the energy change as a bond rotates. The non-bonded interactions $V_{non-bonded}$ are decomposed into a term which accounts for the coulomb interaction between two atoms i and j and a second term which accounts for the attractive van

der Waals interaction (dispersion) and the Pauli repulsion, modeled by the 6-12 Lennard-Jones potential.

$$V_{non-bonded} = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + \frac{q_i q_j}{4\pi\epsilon_0 \epsilon_r r_{ij}} \right) \tag{2.40}$$

ϵ_0 denotes the dielectric constant *in vaccuo* and ϵ_r is the dielectric constant of the medium. The non-bonded interactions are only calculated between atoms in different molecules or, if in the same molecule, separated by at least three bonds.

As in most force fields, the concept of atom types is used in the AMBER force field. An atom type contains not only the information about the element but also information about its hybridization state and sometimes its local environment. So for example, the atom types for single and double bonded carbons would be different. The parameters, such as equilibrium distances and angles, force constants and Lennard-Jones Parameters for each interaction term between atom types are specific for each force field. *Ab initio* quantum chemical calculations or experimental data is usually used as reference data for the force field parametrization.

The partial charges of each atom type are calculated for individual residues in vaccum employing the HF method with a small basis set. This approach yields dipole moments which are 20% too large on average and would not be considered accurate with respect to the results of the recent state-of-the art quantum chemical methods and basis sets. However, this over-polarization allows to implicitly incorporate electronic polarization effects into the force field model.

2.2.2 Molecular Dynamics

The derivative of the force field term specifies the forces acting on each particle i of the molecular system. The dynamics of the molecular system is determined by the newtonian equations of motion for each particle. These N coupled differential equations can not be solved analytically and a finite difference method is used instead in order to numerically integrate the equations of motion. From the force acting on a particle at time t accelerations can be obtained, which together with the velocities and postions at time t are used to calculate the velocities and positions at time $t+\Delta t$. There are several algorithms available which all assume that the positions, velocities and accelerations can be approximated by a Taylor series expansion. The most popular ones are the *Verlet algorithm* [99] and one of its variations the *leap-frog algorithm* [100], the latter used in this work. Both algorithms conserve energy and are time-reversible. The integration time step is determined by the fastest motion in the system. Usually, this is the stretching vibration of a bond involving a hydrogen atom. A time step of 1 fs is often used in practice.

Integration of the equations of motions as described above yields a microcanonical ensemble. Realistic simulations which resemble laboratory conditions need to create a canonical ensemble which conserves not only the total energy but also temperature T and/or pressure P. Constant

temperature is achieved by coupling the system to a bath via a thermostat algorithm. The Nosé-Hoover thermostat [101] ensures the correct ensemble and is realized by introducing the bath as an additional degree of freedom in form of a friction term in the equations of motion. The coupling strength of the thermostat is closely related to period of oscillations between the kinetic energy of the system and the bath. Barostats are implemented in a similar way, allowing the system to fluctuate according to the specified external pressure. The Parinello-Rahman method [102] ensures the correct NPT ensemble.

2.2.3 Free Energy Calculation

The free energies of molecular systems describe their tendencies to associate and react. Free energy calculations are for example used to asses ligand binding energies or to estimate the environmental effects on reactions in proteins and solution.

However, the free energy can not be calculated accurately from a standard MD simulation as described in Sec. 2.2.2. In a standard MD simulation states of low energy are preferentially generated as they represent the equilibrium state. The less probable high energy states are not sampled adequately. The internal energy U for a system of N particles with coordinates $\mathbf{r}^N$, momenta $\mathbf{p}^N$ and particle density $\rho(\mathbf{p}^N, \mathbf{r}^N)$ and energy $E(\mathbf{p}^N, \mathbf{r}^N)$ (normalization factors are omitted for clarity) is definded as

$$U = \int\int d\mathbf{p}^N d\mathbf{p}^r E(\mathbf{p}^N, \mathbf{r}^N)\rho(\mathbf{p}^N, \mathbf{r}^N). \tag{2.41}$$

Eq. 2.41 shows, that the less probable high energy states make an insignificant contribution to the internal energy. Therefore, sampling the lower energy regions of phase space is sufficient in order to accurately calculate the internal energy U.

The free energy of A on the other hand depends exponentially on the energy E

$$A = k_B T ln\left(\int\int d\mathbf{p}^N d\mathbf{p}^r exp\left(\frac{E(\mathbf{p}^N, \mathbf{r}^N)}{k_b T}\right)\rho(\mathbf{p}^N, \mathbf{r}^N)\right). \tag{2.42}$$

Therefore, the high energy states make a significant contribution to the free energy and it is not possible to obtain converged and accurate values from a conventional simulation.

However, when considering a chemical reaction it is sufficient to know the free energy difference between the reactant and product states not the absolute free energies. There are two approaches which make use of this: thermodynamic perturbation or thermodynamic integration [78]. The change of the free energy as a function of a reaction coordinate can be calculated using the potential of mean force (PMF) approach. The reaction coordinate may be for example the distance between a proton and its acceptor in case of a proton transfer reaction or a dihedral angle in case of conformational transitions. The PMF is calculated for a physically achievable process,

consequently yielding a free energy profile in which the point of highest energy corresponds to the transition state of the process[78]. From the energy barrier kinetic quantities such as rate constants can be derived.

The simplest type of PMF is the free energy change as the separation r between two particles is changed. The potential of mean force ω is the representative potential whose gradient gives the averaged force over all configurations of all the 3...N particles acting on particle 1, while keeping particles 1 and 2 fixed:

$$-\frac{d}{d\mathbf{r}_1}\omega^{(2)} = -\langle\frac{d}{d\mathbf{r}_1}V\rangle_{\mathbf{r}_1,\mathbf{r}_2 fixed} = \frac{-\int d\mathbf{r}_3,...,d\mathbf{r}_N(dV/d\mathbf{r}_1)e^{-\beta V}}{\int d\mathbf{r}_3,...,d\mathbf{r}_N e^{-\beta V}} \tag{2.43}$$

with V being the potential which describes the interaction between the N particles. In practice this is obtained by performing an MD simulation in which all degrees of freedom are sampled except for the reaction coordinate $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$. The PMF is directly related to the free energy for changing the system along the reaction coordinate **r** and is also referred to as reversibel work.

In order to sample the unfavorable states sufficiently, a biasing potential $V_i(q)$ is added to the potential energy function

$$V'(\mathbf{r}^N) = V(\mathbf{r}^N) + V_i(q). \tag{2.44}$$

Usually, i biasing potentials are applied for different, restricted intervals of the reaction coordinate and for each potential V_i a separate simulation is carried out. The separate simulations are referred to as "windows" and yield a set of histograms of the biased probability density $\rho_i^\star$. This technique is referred to as Umbrella Sampling. The PMF $\omega(q)$ of the unbiased system can be obtained from the biased probability density $\rho_i^\star$ and the biasing potential $V_i(q)$

$$\omega_i(q) = \Delta A(q) = -k_B T ln\rho_i^\star - V_i(q) + c_i. \tag{2.45}$$

The distributions from each window have to be connected in order to obtain the PMF along the entire reaction coordinate. The offset c_i can be obtained by matching the values for $\omega_i(q)$ at the interfaces between the windows or by the weighted histogramm analysis method (WHAM) [103].

2.3 Theory of Biological Electron Transfer

Electron transfer reactions play a central role in the mechanism underlying biological processes such as photosynthesis, DNA repair or signal transduction as described in Chapter 1. The field of biological electron transfer is one of the most mature fields in biophysics and physical biochemistry, yet there remain areas which are not fully understood [104]. In the following the three mechanism which are currently distinguished in the description of electron transfer in biomolecules

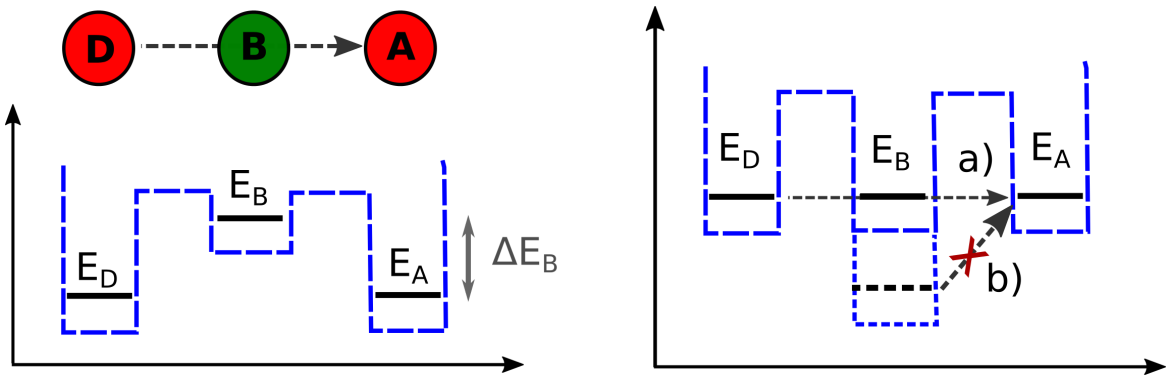


Figure 2.1: Electron transfer regimes. *In the deep tunneling regime the energy of the vibronic bridge (B) states (blue dashed lines) is always higher than the energy of the donor (D) and acceptor (A) vibronic states. If the bridge states can become resonant with the donor and acceptor states (left) resonant tunneling occurs if the bridge residence time is short and (right) hopping takes place if the bridge responds to the charge so that the B state energy is lowered and thermal activation becomes necessary for transfer to A. Inspired by [104].*

will be explained. The important concepts of the most widely used approach for the description of biological electron transfer, classical Marcus theory, will be summarized. Furthermore, the limitations of Marcus' theory will be discussed and a recently developed, direct electron transfer simulation approach will be introduced.

2.3.1 Electron Transfer Mechanisms

Biological electron transfer often occurs over long distances from an initial localized donor to a final localized acceptor via intermediate bridge moieties. The bridge moieties can be any part of protein or DNA and may include cofactor molecules or parts of cofactor molecules. The donor and acceptor can be metal atoms, small organic molecules, aminoacids or DNA bases.

Three different transfer regimes can be distinguished [104], deep tunneling, resonant tunneling and thermally activated hopping. The deep tunneling mechanism applies if the energy difference ΔE_B between localized donor or acceptor states and the localized bridge state is much greater than $k_B T$. As shown in Fig. 2.1(left), the electron here tunnels through the bridge and occupies the intermediate bridge states only with very low probability. This mechanism is mostly found if donor and acceptor are metal moieties whose electronic states are separated from the protein electronic states.

In the case of resonant tunneling (Fig. 2.1(right)) the energy difference between localized donor or acceptor states and the localized bridge state ΔE_B is smaller than $k_B T$. Therefore structural or solvent fluctuations can bring the bridge electronic states in resonance with donor and/or accep-

tor states. The electron occupies the bridge state with high probability but the residence time is too short for the bridge atoms to respond to the electron.

The electron is transferred via thermally activated hopping when the residence time on the bridge is long enough to allow relaxation of the bridge atoms due to the electron. The electron is then trapped in the bridge state(Fig. 2.1(right)) and transfer to the acceptor can take place when thermal fluctuations induce a conformation which brings the occupied bridge and the acceptor state in resonance.

Resonant tunneling or hopping is common if the donor and acceptor states are amino acids or organic molecules whose electronic states are similar to the bridge states. The different transfer regimes are reviewed in detail in [104] and the references therein.

2.3.2 Marcus Theory

Long-range electron transfer reactions in biological systems are usually described by Marcus' theory [79, 80, 105, 106] and its extensions [107–111]. The rate of electron transfer between a donor D and an acceptor A

$$D^+ + A \rightleftharpoons D + A^+ \tag{2.46}$$

can be obtained in the high temperature non-adiabatic limit as

$$k_{ET} = \frac{T_{DA}^2}{\hbar}\sqrt{\frac{\pi}{\lambda k_B T}}\exp\left[-\frac{(\Delta G+\lambda)^2}{4\lambda k_B T}\right]. \tag{2.47}$$

assuming weak coupling between donor and acceptor. T_{DA} is the electronic coupling between donor and acceptor, ΔG is the free energy difference between the initial and final state and represents the thermodynamic driving force for electron transfer. λ is the reorganisation energy and describes the energy which would be required to change the molecular structure of the initial state to the relaxed structure of the final state without the electron transfer taking place. The electron transfer rate depends also on the temperature T, Planck's constant $\hbar$ and the Boltzmann constant k_B.

The reorganization energy is decomposed into an outer and and an inner sphere contribution. The outer sphere reorganization energy λ_o arises from the rearrangement of surrounding molecules (protein and solvent molecules) due to the redistribution of the electron. The inner sphere reorganization energy λ_i represents the change of the molecular structure of donor and acceptor.

The potential energy curves of the initial and final diabatic states are approximated by harmonic parabolas (Fig. 2.2) as a function of a reaction coordinate of the complete molecular system. In case of electronic interaction between the initial and the final state the electron transfer

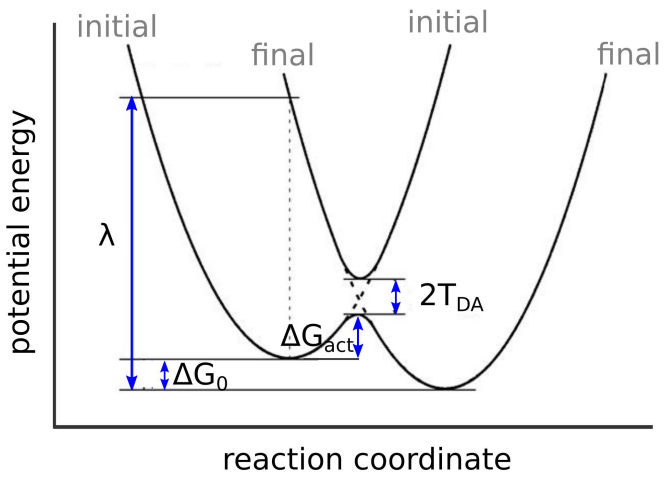


Figure 2.2: Parameters of the Marcus equation. *The potential energy surfaces of the diabatic donor and acceptor states (including dashed lines) are represented by parabolas. Electronic interaction between initial and final state leads to energy splitting of twice the electronic coupling T_{DA}. The reorganization energy λ defines the curvature of the parabolas and ΔG_0 indicates the free energy difference between the inital and final state.*

reaction is adiabatic if the energy splitting $2T_{DA}$ is large or non-adiabatic if the energy splitting is weak $2T_{DA} < k_B T$.

The law of energy conservation and the Franck-Condon Principle require that the electron transfer within Marcus' theory occurs at the crossing of the diabatic states. At the crossing the initial and final electronic states are resonant. Usually the system minimum energy conformation in the initial state is not in a resonance conformation. Thermal fluctuations of the reaction coordinate caused by molecular motion bring initial and final state into resonance and thus enable electron transfer. The activation barrier ΔG_{act} which needs to be overcome to assume the resonance state depends on the reorganization energy λ. Within classical Marcus' theory the reorganization energy is assumed to be the same for forward and backward transfer and the activation barrier can be written as:

$$\Delta G_{act} = \frac{(\Delta G_0 + \lambda)^2}{4\lambda} \tag{2.48}$$

In case of identical species $\Delta G_0 = 0$ and ΔG_{act} becomes $1/4\lambda$.

The Marcus parameters can be computed from simulations. The electronic coupling T_{DA} and the inner sphere reorganization λ_i can be obtained with quantum chemical methods while the free energy ΔG and λ_O are thermodynamic quantities which can be obtained from free energy calculations (Sec 2.2.3).

Several assumptions are inherent to Marcus' theory. Knowledge about the mechanism of transport is required and the electronic structure of the donor and acceptor must be determined beforehand. Also, fluctuation of the tunneling matrix element due to molecular motion is not accounted for in the classical Marcus' equation. Therefore, one recent development in the field of electron transfer simulations is a stronger focus on the coupling between atomic conformational fluctuations and electronic structure changes in biomolecules[112–116]. It is found that energetic and structural fluctuations as well as the relative time scales on which these the fluctuations occur, play a vital role for system properties at the transition state and characterize different electron transfer mechanisms like, e.g. gated electron transfer[112, 113, 117].

Another assumption used to derive Eq. 2.47 is that, prior to the electron transfer, the donor state is thermally fully equilibrated with respect to the conformational dynamics of the system. This assumption is not valid if the timescale for electron transfer is similar to the timescale of relaxation of the reaction coordinate motion. It has for example been shown that electronic and nuclear time scales overlap for the initial electron transfer step in the bacterial reaction center [118, 119]. Here, theoretical and computational descriptions which go beyond Marcus' theory are needed.

2.3.3 Direct Electron Transfer

An alternative to the Marcus model provide non-adiabatic molecular dynamics schemes which propagate the electronic wavefunction and the nuclei simultaneously. The simultaneous propagation allows charge delocalization and takes into account the response of the charge to the nuclear motion as well as the response of the nuclei to the moving charge. The nuclear degrees of freedom are propagated using classical equations of motion and the electronic degrees of freedom are propagated by integrating the time-dependent Schrödinger equation using either a mean-field (Ehrenfest)[120] or surface hopping approach [121, 122]. The electronic structure is computed with DFT or post Hartree-Fock methods. This approach becomes prohibitively costly for large systems involving about 100 QM atoms and electron transfer reactions in the order of nanoseconds.

Recently, an approximative non-adiabatic molecular dynamics scheme has been developed [82, 123–125]. It was shown to reproduces the predictions of high-level methodology and makes electron transfer simulations over long time scales and large systems feasable. This is achieved by splitting the system into a QM and a MM part, describing only the electron and the molecular fragments considered to participate in the transfer quantum mechanically. The rest of the system is treated in terms of classical mechanics. The most time-consuming part are the quantum-chemical calculations which have to be performed in every step. This is accounted for by only taking the frontier orbitals into account (Fragment Molecular Orbital Approach (FMO)) and using semi-empirical DFTB2 for the calculation of the electron transfer parameters. The applica-

tion of the approximative DFT method DFTB [95] reduces the computational time by two to three orders of magnitude compared to conventional DFT, but still maintaining the necessary level of accuracy.

2.3.3.1 Total Energy Derivation

In order to describe the electronic structure of the quantum system we start from a charge-neutral closed shell system with N electrons:

$$E^N[\rho_0] = \sum_i^N \langle \psi_i^0 | \hat{H}[\rho_0] | \psi_i^0 \rangle + E_{DC}[\rho_0] \tag{2.49}$$

where ρ_0 is the DFT ground-state density and ψ_i^0 are the KS orbitals. $E_{DC}[\rho_0]$ contains the DFT 'double-counting' and nucleus-nucleus repulsion terms. Applying the frozen orbital approximation, assuming that $\psi_i^0 = \psi_i^{N-1}$, the energy of the N-1 electron system is obtained by expanding the neutral energy E^N with respect to the density $\rho = \rho_0 + \delta\rho$ around the N-electron density ρ_0 up to the second order.

$$E^{N-1}[\rho] \simeq E^N[\rho_0] - \langle \psi_i^0 | \hat{H}[\rho_0] | \psi_i^0 \rangle + E^{2nd}[\delta\rho] \tag{2.50}$$

This means that the total energy of the ionized system can be expressed as the energy of the neutral system minus the energy of one electron in the orbital ψ_i^0, governed by the Hamiltonian $\hat{H}[\rho_0]$, corrected by electron relaxation effects given through the second order derivative E^{2nd}. The three contributions to the total energy are further approximated as described in the following.

The total energy of the neutral reference system E^N is approximated by the classical molecular-mechanics total energy.

$$E^N[\rho_0] \simeq E^{MM} \tag{2.51}$$

This simple decomposition of the energy into an MM and QM part contributes significantly to the efficiency of the computational scheme.

The wavefunction ψ_i^0 is expanded in the set of HOMO orbitals ϕ_m of the molecular fragments. The fragment orbitals are obtained as DFT (KS) orbitals of the neutral molecular fragments.

$$\psi_i^0 = \sum_{m=1}^{M} a_m^i \phi_m^i \tag{2.52}$$

Computation of the matrix elements H_{mn} is the most time consuming step of the entire electron transfer calculation and the use of fragment orbitals scales this process down to an O(N)-step. In this work, the fragments are the sidechains of the amino acids involved in the electron transfer. As an example the fragment orbitals of three tryptophan side chains in *A. thal.* cryptochrome are

shown in Fig. 2.3. With this expansion the matrix elements

$$\langle \psi_i^0 | H[\rho_0 | \psi_i^0 \rangle = \sum_m a_m a_m^* \langle \phi_n | H[\rho_0 | \phi_m \rangle = \sum_{mn} a_m a_m^* H_{mn}^0 \tag{2.53}$$

are obtained. The diagonal matrix elements $H_{mm}^0 = \varepsilon_m^0$ are the site energies of the neutral fragments and the off-diagonal terms represent the electronic couplings between the HOMOs of the fragments, similar to the parameters in the Marcus equation (Eq. 2.47).

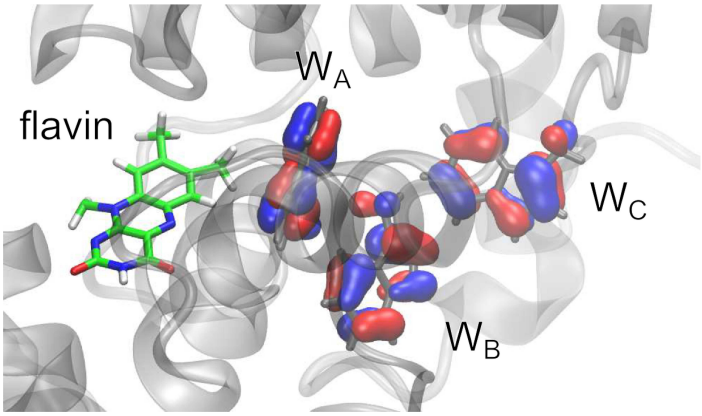


Figure 2.3: Fragment orbital QM scheme. *Shown are the HOMO orbitals of the three tryptophan fragments W_A, W_B and W_C in color, embedded in the classical MM surroundings of the cryptochrome protein (grey), including the flavin cofactor and solvent around the protein.*

Any shift due to ionization is covered by the second order term. The corrected energy for the site energies is obtained by approximating the second order term. The differential density $\delta\rho = \rho - \rho_0$ can also be decomposed into contributions located on the individual molecular fragments

$$\delta\rho = \sum \delta\rho_m \tag{2.54}$$

and the second order term can be written as

$$E_{2nd} = \frac{1}{2} \sum_{nm} \int\int \left(\frac{1}{r - r'} + f_{XC}[\rho_0] \right) \delta\rho_m(r) \delta\rho_n(r') d^3r d^3r' \tag{2.55}$$

Applying a monopole approximation the expression can be further simplified to

$$E_{2nd} = \frac{1}{2} \sum_m U_m \Delta Q_m^2 + \frac{1}{2} \sum_{m \neq n} \frac{\Delta Q_m \Delta Q_n}{R_{mn}} \tag{2.56}$$

where ΔQ_m denotes the fraction of the electron localized on fragment m. The first term contains the Hubbard parameter U (chemical hardness) which determines the electron-electron re-

pulsion on site m. The interaction between neighboring sites is described as simple Coulomb repulsion of the delocalized electron on fragments m and n. This notation is used for simplicity only since the fragment charges Q_m are projected to atomic charges q_α which allows a much more accurate computation of interaction energy. In order to account for the well-known DFT self-interaction error which is incorporated in the second order term a simple self-interaction correction (SIC) is applied by scaling E_{2nd} by a factor $\Gamma = 1/6$. U is determined from DFTB as the derivative of the site energy with respect to the charge on the site using Janaks theorem [90]. A value of $U_W = 5.33eV/e^2$ for the tryptophan side chain is used and $U_Y = 5.86eV/e^2$ for tyrosin.

Now the total energy reads

$$E^{N-1}[\rho] \approx E^{MM} - \sum_{mn} a_m^* a_n H_{mn} + \Gamma\left(\frac{1}{2}\sum_m U_m \Delta Q_m^2 + \frac{1}{2}\sum_{m\neq n}\frac{\Delta Q_m \Delta Q_n}{R_{mn}}\right) \tag{2.57}$$

The site energy of a fragment m in the charged system is obtained as the corrected site energy of a fragment m in the neutral system

$$H_{mm} = \varepsilon_m = -\varepsilon_m^0 + U_m \Delta Q_m + \sum_{n\neq m}\frac{\Delta Q_n}{R_{mn}} \tag{2.58}$$

The influence of the electrostatic potential (ESP) of the complete system at fragment m is accounted for by evaluating the Hamilton matrix elements in the presence of the solvent and the remainder of the protein represented by point charges.

$$ESP_m = \sum_{MMcharges}\frac{q_A^{MM}}{R_{mA}} \tag{2.59}$$

2.3.3.2 Internal Reorganization Energy

The additional Coulomb interaction between fragments due to the presence of the charge is well accounted for by updating the atomic charges of the QM region. The change of geometry within a fragment due to the presence of the charge, called internal reorganization energy, is not included in the equations above. This is due to the fact that the empirical force field parameters are derived for neutral amino acids. In this study, the effect of the internal reorganization energy for each site m is approximated as a correction λ_i^m to the diagonal elements ε_m :

$$\varepsilon' = \varepsilon_m - \lambda_i^m \Delta Q_m^2 \tag{2.60}$$

Values of 0.21 eV and 0.18 eV were obtained for λ_i for tyrosine and tryptophan, respectively. Compared to the SIC scaling parameter Γ, the inclusion of λ_i^m is a minor correction and it may be considered as an additional localization force.

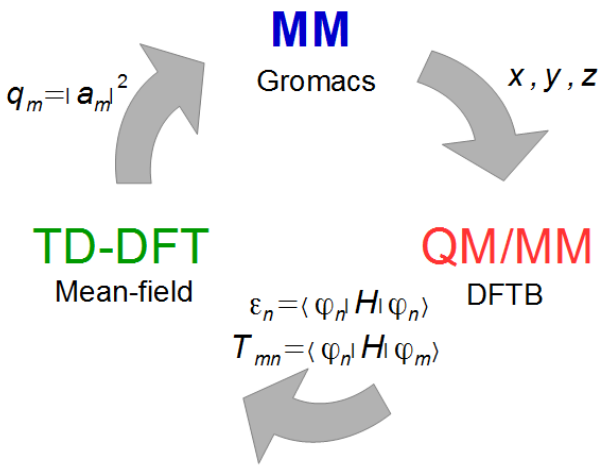


Figure 2.4: Direct electron transfer scheme. *The Hamiltonian from the electronic structure calculation with DFTB is used to propagate the electronic degrees of freedom via TD-DFT. The resulting electron density is projected to atomic charges that are subjected to the next molecular dynamics step. After this, new coordinates for the subsequent electronic structure calculation are obtained.*

2.3.3.3 Propagation of the Charge

Applying the Lagrangian formalism to the total energy (Eq. 2.57), coupled equations of motions for the electronic and nuclear degrees of freedom are obtained [82]. At every MD step the time dependent Schrödinger equation is integrated which leads to propagation of the coefficients a_m

$$i\hbar\frac{\partial a_m}{\partial t} = \sum_n a_n \langle \phi_m | \hat{H} | \phi_n \rangle = \sum_n a_n H_{mn} = \sum_{n \neq m} a_n T_{mn} + a_m \epsilon_m. \tag{2.61}$$

The coefficients a_m are directly related to the charge population on site m at time t via Eq. 2.52

$$Q_m = e|a_m|^2 \tag{2.62}$$

The non-adiabatic simulation of charge transfer described by the theoretical framework above is summarized in 2.4. Since the two sets of equations of motion are coupled, the electronic degrees of freedom are propagated with a variable time step between two time steps of the propagation of the nuclei. The propagation of the electronic degrees of freedom leads to a new electron density distribution on each fragment m. This distribution is projected onto the classical particles by updating the atomic partial charges and used in the following MD step. The propagation of the classical system then proceeds using classical Newtonian mechanics. By using the QM/MM approach the impact of the environment on the traveling electron is considered and by updating

the atomic charges of the charge carriers the polarization of the environment due to the traveling charge is accounted for.

2.3.3.4 Electron Transfer Rates

With the direct electron transfer approach, the time series of the population (Eq. 2.62) can be computed. The time series of the population have been shown to be dependent on the starting structure used [126]. The average population obtained for an ensemble of representative starting structures can be fitted to a kinetic model. The numerical fit yields electron transfer rates which can be compared to experimental data. The model describes a two-step, linear electron transfer between three molecular fragments A, B and C and allows both forward and backward transfer reactions.

$$\mathrm{A} \underset{k_b^I}{\overset{k_f^I}{\leftrightarrows}} \mathrm{B} \underset{k_b^{II}}{\overset{k_f^{II}}{\leftrightarrows}} \mathrm{C} \tag{2.63}$$

k^I and k^{II} denote the rate constants for electron transfer steps I and II and the subscripts f and b indicate the forward and backward transfers, respectively. The two-step kinetic process including back transfer can be described with coupled rate equations:

$$\frac{d[A]}{dt} = -k_f^I[A] + k_b^I[B] \tag{2.64}$$

$$\frac{d[B]}{dt} = -k_f^{II}[B] + k_f^I[A] + k_b^{II}[C] \tag{2.65}$$

$$\frac{d[C]}{dt} = -k_b^{II}[C] + k_f^{II}[B]. \tag{2.66}$$

Here, the square brackets [...] denote the normalized concentration of the corresponding fragment, or in other words, the fragment population. Assuming an electron transfer reaction with initial population of fragment A, the initial conditions for Eqs. (2.64)-(2.66) are

$$[A]|_{t=0} = 1; \qquad [B]|_{t=0} = 0; \qquad [C]|_{t=0} = 0. \tag{2.67}$$

Eqs.(2.64)-(2.66) are solved numerically by varying the four rate constants k_f^I, k_b^I, k_f^{II}, k_b^{II} until the deviation of the kinetic fit from average population is minimized.

CHAPTER 3

Electron Transfer in *E. coli* DNA photolyase: Understanding Polarization and Stabilization Effects via QM/MM Simulations

DNA photolyase facilitates an efficient long-range electron transport between its FAD cofactor and the protein surface via a chain of evolutionary conserved tryptophan residues [44, 45]. The photo-activation is a difficult case for the classical theory of electron transfer. As the electron transfer occurs over nanometer distances and on the timescale of hundreds of picoseconds [40, 41, 44], the use of *ab initio* wavefunction approaches is impossible. Modeling of the photo-activation reaction has therefore mainly been based on classical Marcus' theory [127]. Approximate atomistic QM/MM simulations of electron transfer dynamics were shown to successfully capture the important aspects of the photo-activation reaction with qualitative agreement to experiment [126]. In this chapter several remaining open questions regarding the electron transfer mechanism will be discussed along with methodological improvements of the direct electron transfer approach presented in section 2.3.3.

The electron transfer reaction is completed in less than a nanosecond despite the existence of multiple potential trap states. One goal is to determine how likely it is that additional residues influence the electron transfer during the photolyase photo-activation reaction. By including a larger number of aromatic residues along the electron transfer pathway into the quantum description, the effect of neighboring aromatic amino acids on the electron transfer as well as possible alternate routes can be detected.

The electron transfer along the tryptophan triad is found experimentally to proceed in two subsequent steps, each on the time scale of tens of picoseconds [45]. As the electron transfer occurs on the same time scale as the protein response to the electrostatic changes, it was suggested that the time-scale separation as often presupposed in electron transfer studies might be inappropriate for this system [126]. While solvent relaxation times often seem to be fast compared to the electron transfer, as e.g. in the case of DNA [124, 128], it has recently been shown that the response of coupled hydration and protein conformational dynamics to sudden electrostatic changes of chromophores inside a protein can range from few picoseconds to several hundreds of picoseconds[129]. This is on a comparable time scale to the photo-induced electron transfer in photolyase. Analysis of the residence time of the tryptophan radicals allows to test this hypothesis and to determine the underlying molecular mechanism for efficient transfer.

The framework for the direct electron transfer was derived for non-polarizable empirical force fields. These standard force fields do not account for electronic polarization, an effect which has been shown to strongly influence the properties of electron transfer. A simple, though approximate way to include the effect of electronic polarization in the direct electron transfer approach is by scaling the electrostatic potential. Several scaling factors and their impact on the computed electron transfer parameters and rates will be studied.

Parts of this chapter are published as "'Charge Transfer in *E. coli* DNA Photolyase: Understanding Polarization and Stabilization Effects via QM/MM Simulations"' in the Journal of Physical Chemistry B.

3.1 Simulation Setup

The starting structures for the electron transfer simulations were taken from a 30 ns long classical MD trajectory from a previous study[126]. The MD simulation was performed on the inactive state of photolyase with the flavin being in the oxidized state FAD_{ox}. 29 snapshots were taken at 1 ns time intervals from the MD trajectory to serve as starting structures for the electron transfer dynamics. The selected snapshots represent a typical ensemble of photolyase structural variety, with an average C_α-RMSD value of 0.8 Å compared to their common average structure.

An in-house version of GROMACS 4.6 with modifications was used to carry out the coupled QM/MM calculations during the dynamic electron transfer simulations. The improved Amber99 force field [130, 131] was used for the protein. The FAD cofactor was built from riboflavin and adenosine diphosphate (ADP) models using a force field parametrization developed earlier [130, 132]. Atomic charges for both neutral FAD and negatively charged $FAD^{\bullet-}$ were determined from RESP calculations [133]. Force field parameters for the tryptophan cation radical were obtained

as described before [126]. The temperature was kept at 300 K by applying the Nose-Hoover thermostat [101] while the pressure was controlled by the Parinello-Rahman Barostat [102] at a value of 1 atm. The LINCS algorithm [134] was employed to keep bonds involving hydrogen at fixed length, and a time step of 1 fs was used.

3.2 QM Region Selection

Six residues were chosen to be included in the QM region: the three evolutionary conserved tryptophan residues that make up the established electron transfer chain, W382, W359 and W306 as well as three additional residues, W384, W316 and Y464, close enough to the first three to potentially play a role in the ET reaction. Among the three additional residues, W384 is situated close to both the FAD cofactor and W382 and could be involved in the initial stages of electron transfer. W316 is located close enough to both W382 and W359 to act as a potential fourth electron donor. Experimental evidence of the formation of a tyrosine radical has been found in *A. Nidulans* photolyase [135]. In *E. coli* photolyase Y464 is found close enough to the final W306 that it could function as an alternative terminal electron donor. These six residues will be referred to as electron transfer sites in the following and are shown in Fig. 3.1.

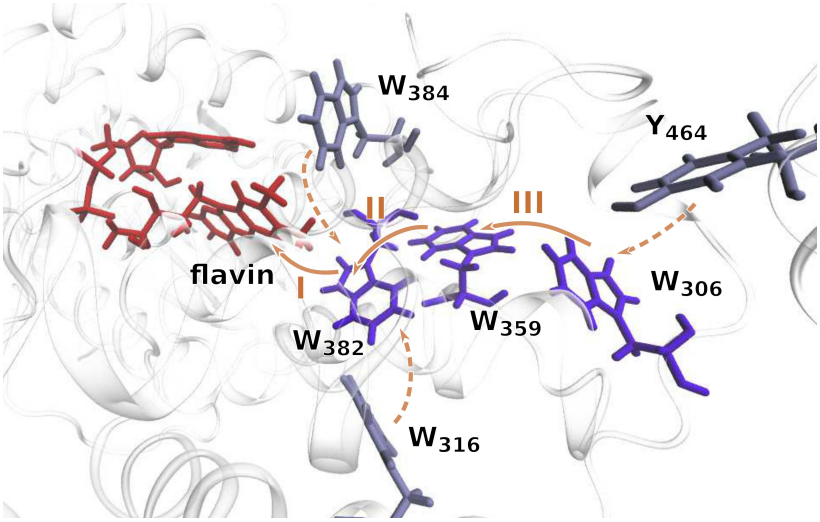


Figure 3.1: Photolyase active site. *During photo-activation an electron is transferred to the flavin cofactor along the tryptophan triad W382, W359 and W306 (violet). The three electron transfers W382→ FAD*, W359 → W382$^{\bullet+}$, W306 → W359$^{\bullet+}$ are labeled I, II and III, respectively. Tryptophan and tyrosine residues in close vicinity to the main pathway (iceblue) might affect the electron transfer.*

The probability for electron transfer in proteins decreases exponentially with increasing dis-

Table 3.1: Distances between all QM region residues *Obtained as average over 1 ns length MD simulations. All distances are given in nm.*

	W384	W316	W359	W306	Y464
W382	1.077	0.735	0.806	1.158	1.574
W384		1.587	1.294	1.634	1.437
W316			1.374	1.374	1.796
W359				0.727	1.282
W306					0.903

tance between donor and acceptor. Thus, the distances between the selected electron transfer sites can indicate potential electron transfer pathways. The average center of mass distance between all six selected residues is summarized in Tab. 3.1.

The average distance between the conserved tryptophans is found to be in the range of 7 to 8 Å. W316 is located at about 7 Å from W382, being the only residue with an electron transfer distance similar to those between the conserved tryptophans. Y464 is found at a distance of about 9 Å from W306. The distances between all other residue pairs amount to more than 10 Å.

3.3 Electron Transfer Dynamics

The photo-activation occurs via subsequent electron transfer steps. In the first step the electron is transferred from W382 to the flavin, yielding a $W382^{\bullet+}$ radical. The missing electron on the $W382^{\bullet+}$ radical is replaced by transfer of an electron from the neighboring W359. Then, the missing electron on the $W359^{\bullet+}$ radical is replaced by transfer of an electron from the neighboring W306. The missing electron can be referred to as a "hole", a positive charge of $+e$, in reference to the nomenclature used in solid state physics. The stepwise electron transfer from W306 to W382 can therefore be described as transfer of a hole from W382 over W359 to W306 as shown in Fig: 3.2. By the direct electron transfer approach, effectively the movement of a positive charge via the HOMO orbitals of the charge carrying residues is simulated. However, here the results will be discussed in terms of electron transfer.

The ensemble of photolyase structure models described above was used as starting point for 29 direct electron transfer simulations. Each individual simulation, independent of the QM region selected and the details of the observed electron transfer process, remained structurally stable with C_α-RMSD values for all 29 QM/MM simulations in the range of 0.8 to 1.5 Å. The initial configuration of the electron transfer simulation is the cation radical $W382^{\bullet+}$ and semireduced flavin $FAD^{\bullet-}$, assuming that the first electron transfer step from W382 to the flavin has been completed. The $FAD^{\bullet-}$ is part of the MM region and remains negatively charged throughout the simulation. The first electron transfer step (see Fig. 3.1) occurs in less than a picosecond after FAD photo-

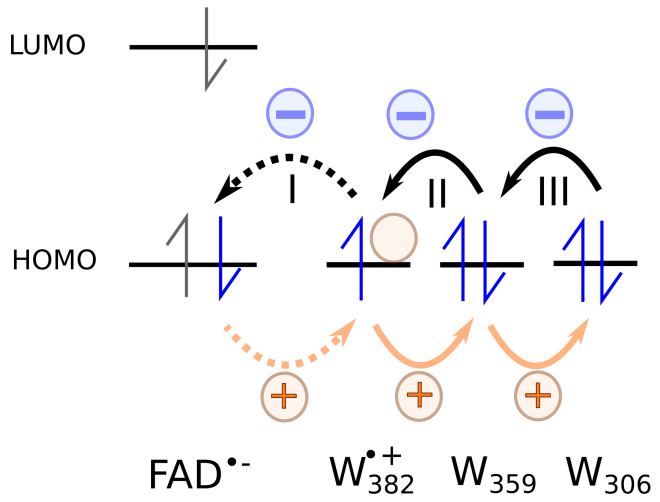


Figure 3.2: Hole transfer during photo-activation of photolyase. *In the first step the electron is transferred from W382 to the flavin, yielding a W382$^{\bullet+}$ and a FAD$^{\bullet-}$ radical. The missing electron on W382 can be referred to as a "hole", a positive charge of $+e$. The stepwise electron transfer from W306 to W382 can therefore be described as transfer of a hole from W382 over W359 to W306.*

excitation [136] and will likely not affect the protein environment before the second step occurs. The electron is free to move between the electron transfer sites of the QM region.

The results of the direct electron transfer dynamics are summarized in Fig. 3.3A, which shows the time evolution of the average population of the electron transfer sites over an interval of 1 ns. The population $Q = |a_m|^2$ (Eq. 2.62) defines which of the electron transfer sites forms the radical at a given time instance. Per definition, the population of an electron transfer site varies between 0 and 1, where 0 is characteristic for a neutral site, while 1 denotes the cation radical state. A population of 0.5 can be interpreted as electron delocalization between two residues.

A similar picture to that of previous simulations [126], where only the three tryptophans of the triad were included in the QM region, emerges. The W382$^{\bullet+}$ radical decays quickly in less than 100 ps, giving rise to transient population of W359$^{\bullet+}$ and formation of the W306$^{\bullet+}$ cation radical. W359$^{\bullet+}$ is only populated by 20% and decays after about 150 ps, giving rise to the formation of W306$^{\bullet+}$. After about 200 ps, W306$^{\bullet+}$ remains stable at 80% for the rest of the simulation. Small population of W316$^{\bullet+}$ is observed during the first 50 ps. Neither W384, which could transiently contribute an electron in the initial stages of transfer, nor the alternative final electron donor Y464 are populated during any of the simulations.

The combined plot (Fig. 3.3B) of the individual trajectories shows that W316$^{\bullet+}$ is partially populated up to 50%, but only for tens of picoseconds and mainly during the first 250 ps of the sim-

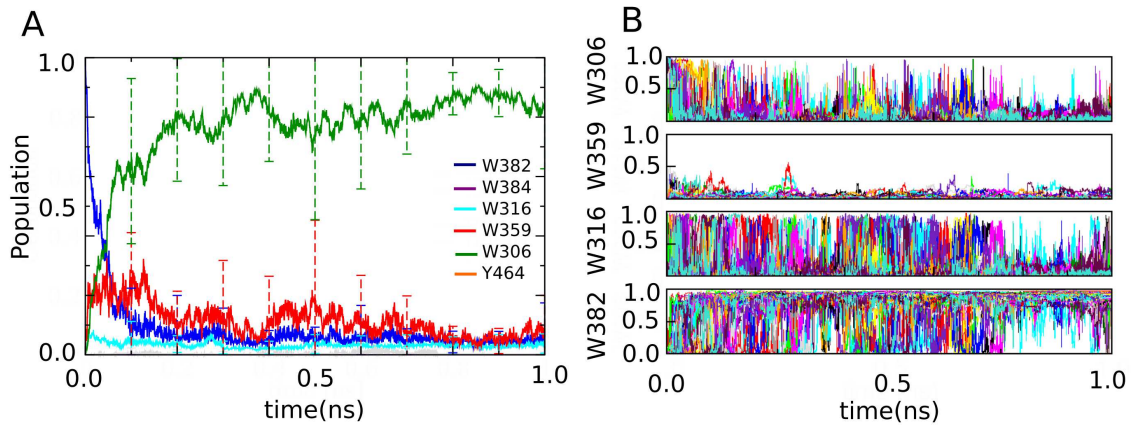


Figure 3.3: Population of the electron transfer sites A *The W382$^{\bullet+}$ decays quickly, giving rise to transient population of W359$^{\bullet+}$ and formation of the W306$^{\bullet+}$ cation radical. The population is obtained as an ensemble average over 29 simulations of 1ns length, which were taken from an MD simulation of the photolyase inactive state. Error bars denote the standard deviation and indicate the variance of* ***B*** *the time evolution observed for the underlying individual simulations. The overlay plot of the time evolution of the individual simulations indicates small, but significant population of the additional electron transfer site W316.*

ulations. For most of the trajectories W316$^{\bullet+}$ is rarely populated by more than 10%. In contrast, W382$^{\bullet+}$, W359$^{\bullet+}$ and W306$^{\bullet+}$ are regularly populated by about 90%. Population on either of the three residues is often only very brief, in the tens of picoseconds range, indicating frequent forward and backward transfer, especially between W382 and W359. After about 800 ps, W306$^{\bullet+}$ is predominantly populated with only occasional, short back-transfers to W359.

The electron transfer is dependent on the time evolution of the ionization potentials and the electronic couplings. The difference of the ionization potentials between two neighboring electron transfer sites is a measure for the dynamical stabilization of a radical by the environment, while the electronic couplings determine the probability for electron transfer between two residues. The average ionization potentials vary significantly, depending on the location of the electron (Figure 3.4). Initially, W382, the starting point for the electron transfer, has the lowest ionization potential, which quickly rises by ca. 1 eV during the initial 100 ps of the simulations. The ionization potentials of W359 and W306 decrease significantly during this time, with the ionization potential of W306 decreasing continuously until it has dropped by ca. 2 eV after 500 ps. This is consistent with the picture of fast initial transfer from W359 to W382$^{\bullet+}$ and further from W306 to W359$^{\bullet+}$, with environmental polarization by the moving electron inducing dynamic charge stabilization and a significant drop in ionization potential.

The additional electron transfer sites W384 and W316, which are located in the protein center,

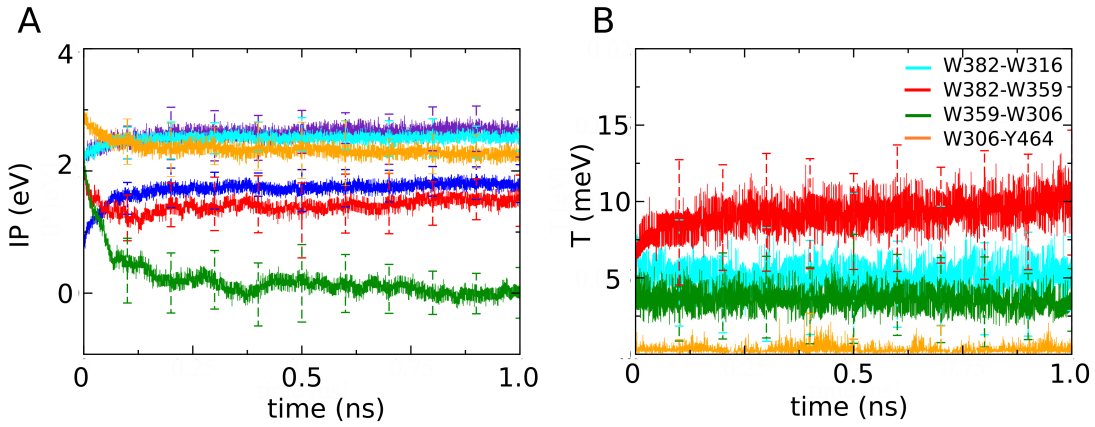


Figure 3.4: Evolution of the electron transfer parameters. A *The ionization potentials of the additional electron transfer sites are energetically separated from the ionization potential of the main sites. The ionization potential of W306 steadily decreases, indicating dynamical stabilization of the W306$^{\bullet+}$ by the environment.* ***B*** *The electronic couplings indicate a high probability for electron transfer only between the main tryptophans as well W382 and W316. Electronic couplings between all other residue pairs are close to zero.*

experience a similar increase in their ionization potential values as W382, while the ionization potential of Y464 drops concurrently with that of W306. Still, the energetic separation between Y464 and W306 remains large, since environmental stabilization of the harder to oxidize tyrosine side chain is not sufficient to allow localization of its radical.

W316 is found to be only coupled to W382, with the electronic coupling decaying from about 7.5 meV in the first picoseconds to a value of about 4.98 ± 0.68 meV for the rest of the simulation. The relatively high electronic coupling during the first picoseconds of the simulation might explain the small observed population of W316. The electronic coupling between W382 and W359 steadily rises from ca. 7 meV at the start of the simulation to ca. 10 meV after 1 ns. The electronic coupling between W359 and W306 is fluctuating around an average value of 3.52 ± 0.51 meV. The time evolution of the electronic couplings shows that there is no coupling between W384 and any of the other residues. The electronic coupling between Y464 and W306 is 0.26 ± 0.17 meV and seems too small to mediate electron transfer between these residues.

3.4 Electron Transfer Rates

To connect the population trajectories to experimental data, we have conducted a numerical fit to two different kinetic models. The kinetic model described in chapter 2.3.3.4, named the three-side model, was extended to include W316. The extended model includes the two reactions described in the three-side model

Table 3.2: Rate constants of the electron transfer steps in E. coli photolyase. *Different kinetic models are used. Error bars indicate the deviation of optimization based on different initial values. If no error bar is given, all optimizations converged to the same value.*

e^- transfer step	donor acceptor	k_f [ns^{-1}]	k_b [ns^{-1}]
fit to three side model			
II:	$W_{359} \rightleftarrows W_{382}$	34.8	18.3
III:	$W_{306} \rightleftarrows W_{359}$	43.3	7.4
fit to extended model			
II:	$W_{359} \rightleftarrows W_{382}$	32.4	5.0±3.3
III:	$W_{306} \rightleftarrows W_{359}$	47.6	7.7
IV:	$W_{316} \rightleftarrows W_{382}$	8.9±1.8	72.4±21.1
V:	$W_{359} \rightleftarrows W_{316}$	0.0	9.8±3.4

$$\mathrm{W382} \underset{k_b^{II}}{\overset{k_f^{II}}{\leftrightarrows}} \mathrm{W359} \underset{k_b^{III}}{\overset{k_f^{III}}{\leftrightarrows}} \mathrm{W306} \tag{3.1}$$

as well as the two additional reactions

$$\mathrm{W382} \underset{k_b^{IV}}{\overset{k_f^{IV}}{\leftrightarrows}} \mathrm{W316} \underset{k_b^{V}}{\overset{k_f^{V}}{\leftrightarrows}} \mathrm{W359}. \tag{3.2}$$

The population of $W384^{\bullet+}$ and $Y464^{\bullet+}$ is negligible over the course of the 1 ns simulation and therefore has been omitted from the kinetic model. Fits to the original and the extended model were conducted as a series of 100 optimizations of 25,000 steps each, starting from random initial values for the rate constants. Error estimates are calculated from the spread in the results of different optimizations. The rate constants obtained are given in Table 3.2. The first fit (top block in Table 3.2) is based on the three side model and disregards the involvement of W316. The rates for the electron transfers $W_{359} \rightarrow W_{382}$ and $W_{306} \rightarrow W_{359}$ are found to be about 35 ns^{-1} and 43 ns^{-1}, respectively. A smaller transfer rate for step II compared to step III as well as small but significant backtransfer rates were also obtained for a fit to the three side model based on simulations using a different set of starting structures and a different QM zone selection [126].

The rates for the transfers $W_{359} \rightarrow W_{382}^{\bullet+}$ and $W_{306} \rightarrow W_{359}^{\bullet+}$ computed with the extended model (second block in Table 3.2) are similar compared to the rates computed with the original three side model only taking into account the transfer between W382, W359 and W306. Only the backtransfer rate $W_{359} \rightarrow W_{382}^{\bullet+}$ is found to be reduced, being three times smaller when computed with the extended model compared to the three site model. The extended model yields a rate of about 9 ns^{-1} for electron transfer between $W_{316} \rightarrow W_{382}^{\bullet+}$ which is about for times smaller than the rates for the transfer steps II and III along the conserved tryptophans. The rate of back electron transfer $W_{316} \rightarrow W_{382}^{\bullet+}$ is much larger than the forward transfer rate and indicates that W316 only donates a

small fraction of charge. No transfer between $W_{359} \rightarrow W^{\bullet+}_{316}$ is found.

3.5 Cause for Efficient Transfer across W359

The ionization potential describes the amount of energy required to remove an electron from a molecule. The ionization potential of a molecule in a protein or solvent environment will differ from its ionization potential in the gas phase due to electrostatic interaction with the environment. The lower the ionization potential of an electron transfer site, the more favorable is the stabilization of the respective radical by the environment.

In order to quantify how well the individual radicals can be stabilized by the environment, separate classical molecular dynamics simulations of each radical state were conducted. A radical state represents the situation in which the electron has been fully transferred from one electron transfer site to another, creating a tryptophan or tyrosine radical. Therefore, in each classical simulation exactly one radical state remained fixed, by using atomic charges derived for the cation radical of the respective residue. The ionization potentials of the electron transfer sites for each radical state of photolyase are summarized in Fig. 3.5. The ionization potentials for each state are obtained as average over 20 ps simulation.

For the inactive state of photolyase the ionization potentials of all the electron transfer sites computed in the protein and solvent environment are similar. This indicates that the protein electrostatic potential does not provide a potential energy surface, brought about, e.g., by a specific arrangement of charged amino acids, that would energetically favor the electron transfer. As found previously, it is rather dynamic re-polarization of the environment by the moving electron which creates the driving force for electron transfer [126].

Comparison of the ionization potentials of electron transfer sites in the different radical states shows that each of the selected tryptophan residues (except of W316) could in principle form a stable radical cation within 20 ps, indicated by the lower ionization potential of about 1 eV of the respective radical relative to all other sites. Among the residues which constitute the tryptophan triad, the $W306^{\bullet+}$ radical shows the lowest ionization potential of about -0.3 eV compared to 0.4 eV for the $W359^{\bullet+}$ radical state and 0.7 eV for the $W382^{\bullet+}$ radical state. The difference between the ionization potentials of these radicals represents the driving force for the electron transfer along the tryptophan triad. The ionization potentials of $W384^{\bullet+}$ and $Y464^{\bullet+}$ are similar to the ionization potential of $W382^{\bullet+}$ and therefore are energetically highly favorable to be populated during the electron transfer. The $W316^{\bullet+}$ radical experiences the least stabilization. $W306^{\bullet+}$ and $Y464^{\bullet+}$ which are located close to the protein surface experience the strongest stabilization compared to the other charge carrying residues.

However, the average population plot obtained from the direct electron transfer dynamics indicates that only the $W382^{\bullet+}$ radical and the $W306^{\bullet+}$ radical are fully stabilized between the electron

transfer steps. That none of the other radicals are found to be fully stabilized during the electron transfer dynamics simulations can be partially rationalized by the electron transfer parameters.

Vanishing electronic coupling prevents the $W384^{\bullet+}$ radical from becoming populated at all. The relatively high electronic coupling between W316 and W382 makes electron transfer between the two sites highly probable. However, backward transfer is not hindered by dynamic stabilization of the $W316^{\bullet+}$ radical as the difference of its ionization potential compared to the ionization potentials of the other sites by about 0.3 eV provides only little environmental stabilization. The $W306^{\bullet+}$ radical in comparison, e.g., which is fully stabilized during the electron transfer dynamics simulation is found to have a lower ionization potential of about 1 eV compared to all other radicals.

The full stabilization of the $W359^{\bullet+}$ radical after 20 ps seems inconsistent with the observed fast electron transfer from W306 to $W359^{\bullet+}$ in the dynamic simulations above. Also, the ionization potential is never found to be minimized during the electron transfer dynamics(see Fig.3.4), indicating that $W359^{\bullet+}$ is never fully stabilized during the dynamics. Previous simulations suggested that the electron transfer from W306 to $W359^{\bullet+}$ occurs faster than polarization of the environment by the $W359^{\bullet+}$ radical[126], so that $W359^{\bullet+}$ radical is only transiently populated and never triggers the full environmental response. In order to investigate this further, the residence time of the $W359^{\bullet+}$ radical during the electron transfer dynamics was computed.

Figure 3.6 shows the residence time of the $W359^{\bullet+}$ radical obtained from the electron transfer dynamics simulation. The residence time is defined by the time interval in which the $W359^{\bullet+}$ radical is populated by more than 60% before electron transfer from W382 or W306 occurs. The time intervals are clustered and it is counted how often the $W359^{\bullet+}$ was found populated in a respective time interval during the electron transfer simulation. The residence time of the $W359^{\bullet+}$ radical is below 1 ps in most cases and never larger than 100 ps. This indicates that during the electron transfer dynamics the $W359^{\bullet+}$ radical is never populated long enough to trigger the full environmental response. As a consequence the electron transfer from W306 to $W359^{+}$ occurs out of a non-equilibrium conformation of the environment.

The very large number of extremely short populations of the $W359^{\bullet+}$ radical in the 10 fs range is caused by rapid fluctuations of the population trajectories. These very short populations occur clustered in rare parts of the trajectories, disappear when curves are smoothed and can be considered artifacts from the histogram generation. Figure 3.6B shows which of the three main radicals constituting the tryptophan triad, $W382^{\bullet+}$, $W359^{\bullet+}$ and $W306^{\bullet+}$, is formed over the course of the simulation for three typical electron transfer trajectories. If a residue is found populated by more than 60% percent, the radical is considered to be formed. The $W359^{\bullet+}$ radical occurs only briefly compared to the longer residence times of 100 ps or more, found for the other two main electron transfer sites, $W382^{\bullet+}$ and $W306^{\bullet+}$. This is confirmed by the histograms showing the residence time of the $W382^{\bullet+}$ and $W382^{\bullet+}$ radicals (Fig.3.7). In addition to low residence times of less than

10 ps, which occur due to structural fluctuations, long residence times of more than 100 ps are observed to a significant extend. This indicates that the $W382^{\bullet+}$ and $W382^{\bullet+}$ radicals are indeed fully stabilized by the environment during the electron transfer dynamics.

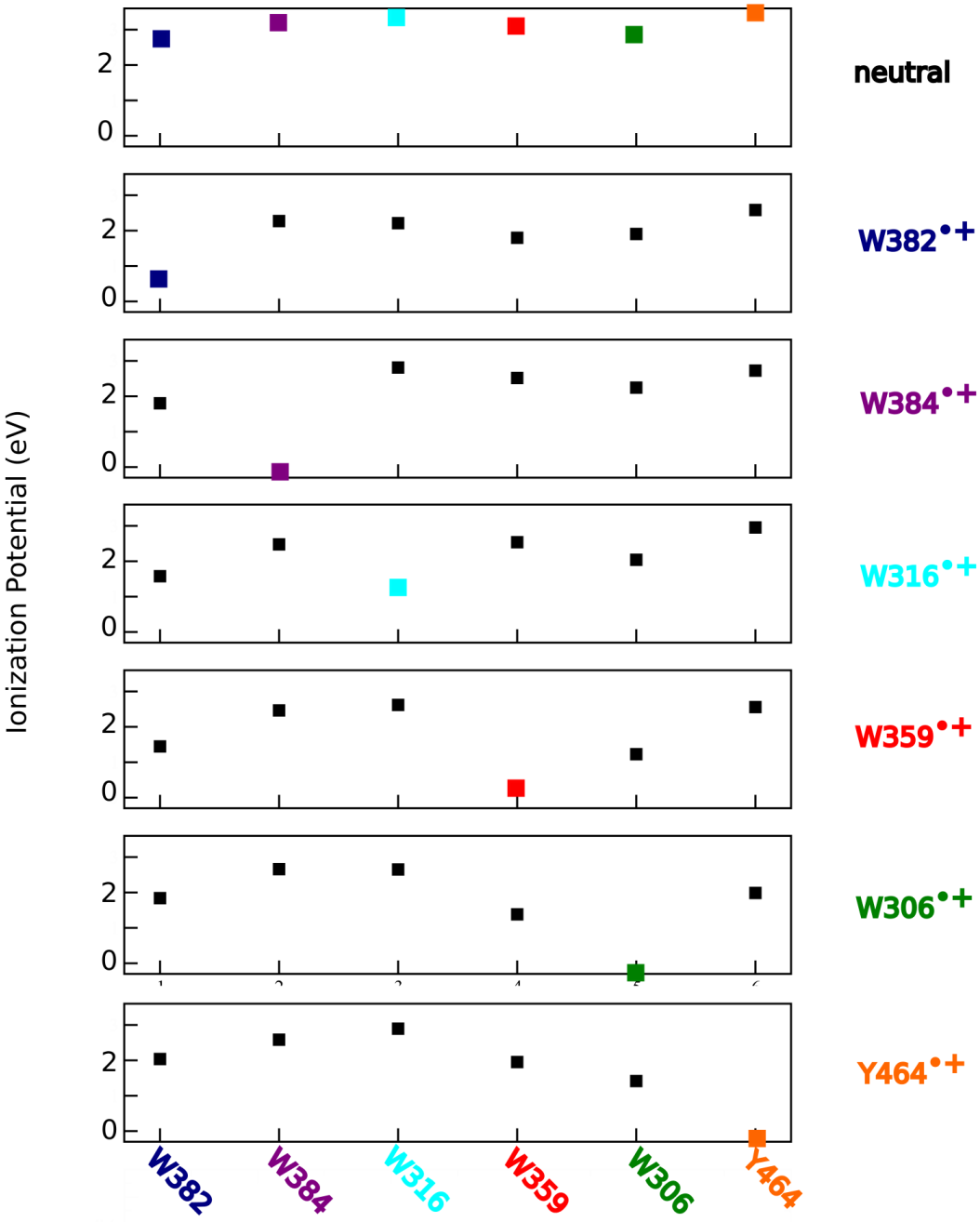


Figure 3.5: Ionization potentials of the electron transfer sites for different radical states. *The charge carrying site always has the lowest potential energy for each radical state, indicating that each electron transfer site could in principle form a stable radical cation within 20 ps. In the photolyase inactive state the ionization potentials are similar for all electron transfer sites, indicating that the protein does not provide a potential energy surface which energetically favors electron transfer. The panels shows the ionization potentials obtained from separate classical MD simulations of the neutral and each radical state.*

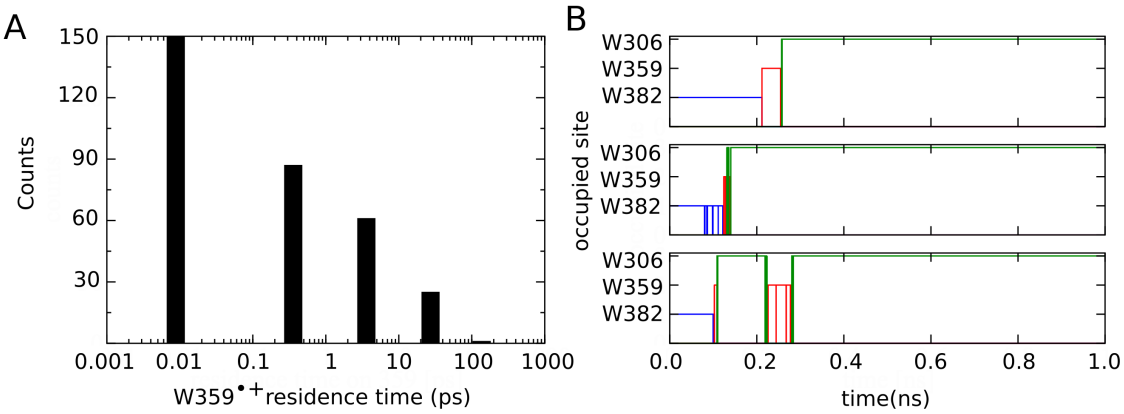


Figure 3.6: Residence time of the W359$^{\bullet+}$ during the electron transfer dynamics.A *The logarithmic histogram shows that the W359$^{\bullet+}$ is mostly populated for less than 10 ps before electron transfer from the neighboring electron transfer sites occurs.* ***B*** *Three examples of the time evolution of radical formation show that the population of W359$^{\bullet+}$ (red) is brief compared to W382$^{\bullet+}$ (blue) and W306$^{\bullet+}$ (green).*

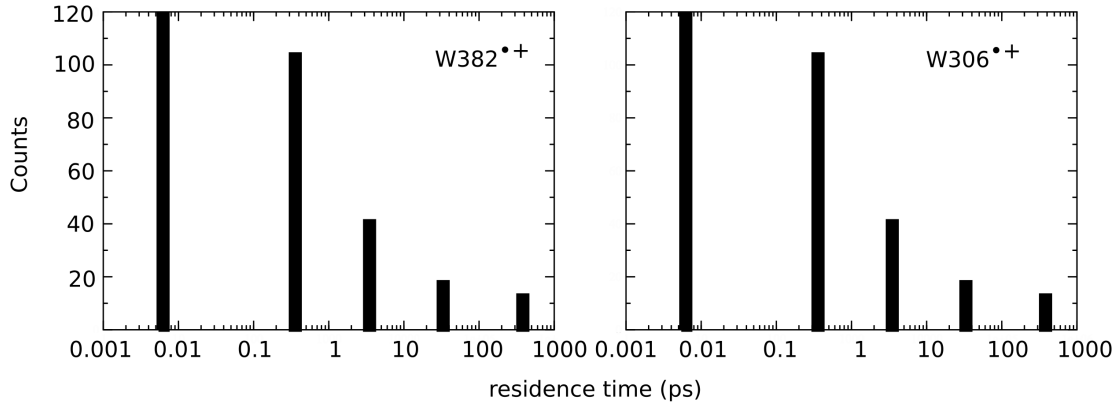


Figure 3.7: Residence time of W382$^{\bullet+}$ and W382$^{\bullet+}$ during the electron transfer dynamics. *The logarithmic histogram shows that the W382$^{\bullet+}$ and W382$^{\bullet+}$ are significantly populated during intervals longer than 100 ps before electron transfer from the neighboring electron transfer sites occurs, indicating full stabilization of the two radicals by the environment.*

3.6 Electrostatic Potential Scaling

The electronic charge density of the molecular environment is represented by fixed point charges using a non-polarizable force field. Non-polarizable force fields neglect the effect of electronic polarization. The effect of electronic polarization, however, has been shown to strongly impact computed charge transfer properties as well as other molecular properties [81, 137–140]. In particularly the reorganization energy λ has been found to be affected by polarization effects. The outer-sphere reorganization energy λ_O can be overestimated by 30-50% when computed with a fixed charge model and has been suggested to be scaled by the inverse optical dielectric constant $1/\epsilon^{opt}$ of the medium [140, 141]. Scaling factors in the range of $1/2 - 1/1.4$ have been used in the past [138, 139, 141]. In the direct electron transfer approach the effect of electronic polarization can be included by scaling the electrostatic potential which is applied during the QM/MM calculations of the matrix elements H_{mn} (see Eq. 2.59)

$$ESP_m = \frac{1}{\epsilon_{MM}} \sum_{MMcharges} \frac{q_A^{MM}}{R_{mA}} \tag{3.3}$$

Effectively, the electrostatic potential resulting from the MM environment as well as the moving electron are scaled down by ϵ_{MM}.

To explore the effect of electronic polarization on the electron transfer dynamics, we have conducted a series of 22 the electron transfer dynamic simulations using different scaling factors ϵ_{MM} of 1.0, 1.4 and 2.0. The QM zone was restricted to W382, W359 and W306 to increase computational efficiency.

Figure 3.8 summarizes the effect of the scaling factor on the the average population of the three charge carrying residues. The average population is markedly affected by changes in ϵ_{MM}. Increasing the scaling factor results in faster electron transfer: employing a scaling factor of $\epsilon_{MM} = 1.0$ yields a persistent $W306^{\bullet+}$ radical after about 300 ps, while a persistent $W306^{\bullet+}$ radical is already obtained after tens of picoseconds if a scaling factor of $\epsilon_{MM} = 2.0$ is used. Also, only about 75% of the $W306^{\bullet+}$ radical is stabilized when a scaling factor of $\epsilon_{MM} = 2.0$ is employed. Additionally, fluctuations in the population increases with higher ϵ_{MM}, most notably for $\epsilon_{MM} = 2.0$. The rate constants obtained from the simulations with varying scaling factors are listed in Table 3.3. It is found that the rate constants for the electron transfer steps increase when ϵ_{MM} is changed from 1.0 to 1.4 and increase slightly further when ϵ_{MM} is set to 2.0.

For the forward transfer steps, much larger rate constants are found in both cases (compare third and fourth block of rates in Table 3.3 to the first block). ϵ_{MM}=1.4 leads to approximately tripling of the rate constant for the transfer $W_{359} \rightarrow W_{382}$ and doubling of the rate constant for the transfer $W_{306} \rightarrow W_{359}$, while ϵ_{MM}=2.0 gives a sixfold increase and a fourfold one for the two respective transfers. The rate constants for the backward reactions also rise significantly.

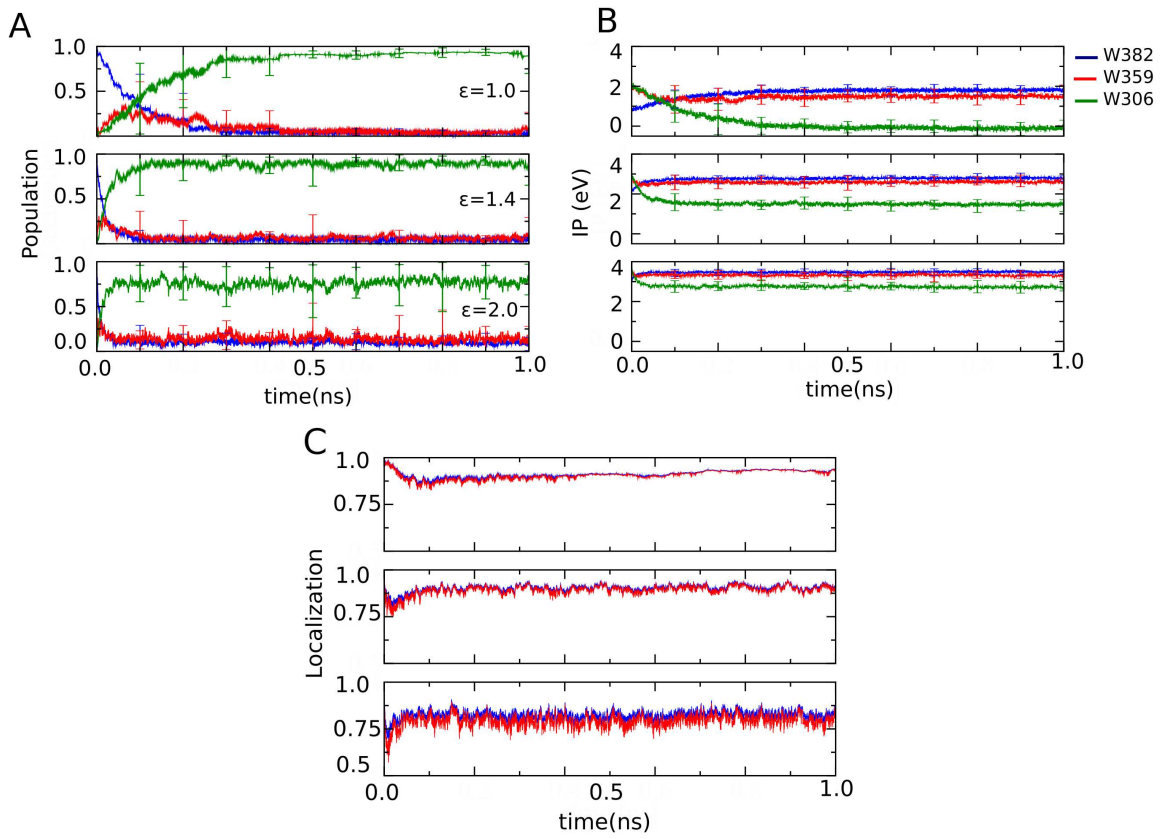


Figure 3.8: Characteristics of electron transfer using different polarization parameters. ***A*** *Electrostatic potential scaling results in faster electron transfer and increases the fluctuation of the average site population.* ***B*** *The energetic separation and fluctuations of the ionization potentials of the electron transfer sites decreases due to scaling of the electrostatic potential,* ***C*** *leading to increased delocalization of the electron. Different colors correspond to two different definitions of the localization parameter.*

The time evolution of the average ionization potentials of the charge carrying residues is also a key quantity to capture the effect of electronic polarization on the electron transfer properties. As shown in chapter 3.5 above, the exergonicity of the electron transfer reaction is mainly established by dynamic re-polarization of the environment and the fact that the $W382^{\bullet+}$ radical close to the protein surface is better stabilized than the $W382^{\bullet+}$ radical inside the protein. In particular, the polarizability of the solvent drives the electron transfer, when the solvent electrostatic potential is excluded, the photo-activation becomes an endothermic, uphill process [126]. The stabilization of an electron transfer site is quantified by its relative ionization potential. The lower the ionization potential compared to the ionization potential of the other electron transfer sites, the better the stabilization of the respective radical by the environment.

Figure 3.8 summarizes the effect of the scaling factors on the ionization potential. With increas-

Table 3.3: *Rate constants of the electron transfer steps in E. coli photolyase for simulations using different polarization parameters. All rates are given in ns^{-1}.*

e^- transfer step	donor acceptor	k_f [ns^{-1}]	k_b [ns^{-1}]
fit for scaling factor ϵ_{MM}=1.0			
II:	$W_{359} \rightleftarrows W_{382}$	34.8	18.3
III:	$W_{306} \rightleftarrows W_{359}$	43.3	7.4
fit for scaling factor ϵ_{MM}=1.4			
II:	$W_{359} \rightleftarrows W_{382}$	88.3	86.7
III:	$W_{306} \rightleftarrows W_{359}$	64.3	6.3
fit for scaling factor ϵ_{MM}=2.0			
II:	$W_{359} \rightleftarrows W_{382}$	179.7	125.4
III:	$W_{306} \rightleftarrows W_{359}$	162.5	29.6

ing scaling factors, the difference between the ionization potentials of the electron transfer sites is reduced. Thus, both the driving force for electron transfer ΔG_0 as well as the reorganization energy lambda λ_O are reduced. Therefore, the scaling of the MM electric field on the one hand makes the reaction less exergonic and and on the other hand lowers the reaction barrier. These two competing effects combined lead to the increase of the rate constants as shown in Fig. 3.8. The reduced driving force also lowers the barrier for back transfer. This implies that the electron is less stabilized during the dynamics and explains the observed faster transfers and the increased number of backtransfers when a scaling factor is employed.

The effect of the scaling factors on the localization properties of the electron is shown in Fig. 3.8C. Two different definitions were used to capture the effect of electron localization. Localization is defined either as the square of the largest orbital expansion coefficient or based on the root of the sum of all squared fractional charges. Both definitions yield very similar results, see Figure 3.8C. For ϵ_{MM} = 1.0, the electron remains localized to more than 85 % over the course of the simulation. For ϵ_{MM} = 1.4,the localization parameters drop slightly, and for ϵ_{MM} = 2.0, only ca. 75 % of the electron is localized on one site. Notably, in all three cases, the localization quickly decreases during the initial stages of the electron transfer and increases once the electron is localized on W359, yielding $W306^{\bullet+}$ as the stable radical.

3.7 Discussion

Previous theoretical studies of the photo-activation reaction in *E. coli* photolyase were compared to experimental data from UV-vis spectroscopy of semiquinone photolyase, estimating the rate constants to be >110 ns^{-1} and >33 ns^{-1} for the respective electron transfer steps[37]. Recently it was found that the oxidation state of the flavin cofactor seems to affect the photo-activation

dynamics and efficiency. Time resolved femtosecond spectroscopy experiments found the first two electron transfer steps in semiquinone photolyase to be significantly slower than in oxidized photolyase [46, 47]. Here, the electron transfer dynamics simulations were carried out for oxidized photolyase. Therefore, the rate constants computed in this study should be discussed with respect to the experimental data from femtosecond spectroscopy experiments of oxidized photolyase [46].

The calculations with an extended QM zone, including six electron transfer sites still yield an essentially unchanged picture of electron transfer in photolyase: the electron subsequently moves from W359 to $W382^{\bullet+}$ and from W306 to $W359^{\bullet+}$, with forward rate constants in the 30 to 40 ns^{-1} range. W316 is found to contribute to the electron flow from the protein surface to the flavin cofactor but to a lesser extend than the conserved tryptophans. The $W316^{\bullet+}$ radical only exists temporarily, the electron donation from W316 to $W382^{\bullet+}$ occurs partially or temporarily with a rate constant of about 7 to 11 ns^{-1} and a much higher rate constant of about 50 to 90 ns^{-1} for the respective backtransfer.

The Ehrenfest propagation scheme used in the simulations is known to have the tendency to overly delocalize the electron. This effect is likely more pronounced if the differences of the ionization potentials between electron transfer sites are small and increases with simulation time. As the difference of less than 0.5 eV between ionization potential of $W382^{\bullet+}$ and that of $W316^{\bullet+}$ is small, the partial electron donation from W316 can not be completely ruled out to be a simulation artifact. However, the fact that small, but significant population of $W316^{\bullet+}$ is mostly observed during the first 100 ps of the simulation opposes that $W316^{\bullet+}$ is found populated in the electron transfer dynamics simulation purely due to artificial electron over-delocalization.

Clarity regarding the role of W316 and qualitative confirmation of our model is provided by the results of the time resolved femtosecond spectroscopy experiment of oxidized photolyase: involvement of W316 in the electron dynamics during the photo-activation reaction has been observed [46]. The experimental kinetics also found busy forward and backward transfer between W316 and W382 with time constants of 12 ps and 40 ps, respectively. These time constants correspond to rate constants of about 25 ns^{-1} for the forward and 83 ns^{-1} for backward transfer and agree well with the rate constants obtained from the direct electron transfer simulation.

For the forward electron transfers $W_{359} \rightarrow W_{382}$ and $W_{306} \rightarrow W_{359}$ time resolved femtosecond spectroscopy of oxidized photolyase yielded time constants of 70 ps and 150 ps which correspond to rate constants of about 14 ns^{-1} and 6 ns^{-1}, respectively, suggesting that the second electron transfer step occurs about twice as fast as the third step. The rates constants for the forward electron transfers $W_{359} \rightarrow W_{382}$ and $W_{306} \rightarrow W_{359}$ computed with the kinetic model are in qualitative agreement with the experimental rates. Rate constants of about 32 $ns^{-1}$and 48 ns^{-1} are obtained which are the same order of magnitude as the experimental rate constants but do not capture the effect of a faster transfer followed by a slower transfer.

Formation of $W384^{\bullet+}$ by direct electron transfer to the $FADH^{\bullet}$ was observed with a time constant of 70 ps (corresponding to a rate constant of about 13 ns^{-1}) during the experiment [46] but no electron transfer between W384 and any of the other tryptophans was found. Direct electron transfer from $W384^{\bullet+}$ to the flavin could not be probed by our model. However, the possibility of the formation of the stable trap state by the $W384^{\bullet+}$ radical was indicated by its low ionization potential compared to the ionization potential of the other electron transfer site in the $W384^{\bullet+}$ radical state.

As discussed above, non-polarizable force fields tend to overestimate the solvent reorganization energy λ_s by roughly a factor 1.4^{-1}. It has been proposed that charged residues or ions in molecular mechanics models should be scaled down by roughly this factor to cover polarization implicitly.[142] The introduction of a scaling factor of $\epsilon_{MM} = 1.4$ is completely in line with this reasoning. Here, polarization is included by scaling down the electrostatic interaction of the electron with the MM environment and leads, in a Marcus picture, to a lower solvent reorganization energy λ_S. This in turn accelerates the electron transfer rates by a factor of 3 to 6, depending on ϵ_{MM}, in our simulations. Importantly, comparison to the experimental data from femtosecond spectroscopy shows that with a polarization scaling parameter of $\epsilon_{MM} = 1.4$ the experimental characteristics are well reproduced, with the rate constant for the second electron transfer step being higher than for the third step.

Importantly, the non-equilibrium character of the electron transfer reaction is captured correctly by the direct electron transfer approach. Straightforward application of the Marcus formula yields a rate constant for the third electron transfer step which is underestimated by several orders of magnitude compared to the experimental values [126]. The reason for this underestimation is, that standard computation of the Marcus parameters ΔG_0 and λ_S involve a complete relaxation of the environment. As shown above, this is not the case in photolyase. $W359^{\bullet+}$ is only stabilized briefly during the electron transfer dynamics, even though its ionization potential is low enough compared to the ionization potential of the other electron transfer sites so that it could serve as an intermediate trap state.

This discrepancy can be understood when taking the delayed environmental response into account. As soon as a positively charged radical is formed, it will polarize its environment which in turn stabilizes the radical. Some of these stabilization processes like solvent reorientation occur fast compared to the time scale of electron transfer and extremely fast in the case of electronic polarization. But in larger biomolecules, much slower relaxation time scales of up to several nanoseconds play important roles as well [128, 129].

Here, it has been shown that a large amount of stabilization is completed after 20 ps. In previous simulations it has been found that longer simulations of the radical states allow additional

slow protein conformational equilibration and increase the energy gap even further (as shown for the case of three QM sites in Figure 12d in Ref. 126). This slower contribution to relaxation does not influence the dynamics for the case of the $W359^{\bullet+}$ radical, which forms only for a few ps. The electron transfer from W306 to $W359^{\bullet+}$ therefore typically occurs out of non-equilibrium conformations and needs to overcome a significantly smaller energy gap than equilibrium simulations would suggest, resulting in much faster transfer rates. Presumably, the structure and dynamics of the photolyase enzyme have evolved to prevent W359 from becoming a charge trap state, thereby leading to a fast and efficient photo-activation process.

This shows the direct electron transfer approach to be a valuable predictive tool to establish transfer pathways in proteins where conclusively experimental validation is lacking, such as other members of the photolyase/cryptochrome protein family. The main electron transfer pathway was identified as well as the small, but significant contribution to the electron flow by the additional W316 residue. Comparison of the computed forward electron transfer rates to experimental data shows that the correct order of magnitude is reproduced. A direct comparison of the back electron transfer rates between theory and experiment is made more difficult since our kinetic modeling scheme required the inclusion of significant very fast rates for immediate backward electron transfer reactions for a good fit, which are not accessible experimentally. After inclusion of electronic polarization by scaling the molecular mechanics electrostatic potential, the ratio of the electron transfer rates for the second and third steps was obtained in accord to the experimental observations. Importantly, non-equilibrium effects which seem to be crucial for the efficiency of the photo-activation reaction are captured correctly by the direct electron transfer approach.

CHAPTER 4

QM/MM Simulation of Electron and Proton Transfer during Photo-activation of *A. thaliana* cryptochrome

In plant cryptochromes from *A. thal.* [64, 143], FAD photo-excitation by blue light leads to conversion of fully oxidized FAD to a semireduced $FADH^{\bullet}$ radical form which represents the signaling state [61, 62]. The conversion happens in due to light-induced electron transfer involving FAD and a chain of three tryptophan amino acid residues W400, W377 and W324 [63, 64, 144], which is found conserved throughout the entire cryptochrome/photolyase family [72]. Triggered by the electron transfer event, a proton transfer from D396 to FAD was reported for *A. thal.* cryptochrome [37, 63, 64] which can further stabilize the ionic radical pair $FAD^{\bullet -}$+$W(H)^{\bullet +}$.

So far, theoretical investigation of the electron transfer focused on the static configurations of cryptochrome taken from MD simulations of different cryptochrome radical pair states. The active site of the protein was studied quantum mechanically in its different possible electronic states. It was demonstrated that a radical pair involving flavin and a tryptophan residue $W_A(H)^{\bullet +}$, see Fig. 4.1, is rapidly formed upon flavin photo-excitation, and could subsequently be stabilized through a proton transfer between positively charged W_A and negatively charged flavin involving aspartic acid as an intermediate, see Fig. 4.1. Further analysis of the electronic structure of the cryptochrome active site [147] demonstrated the possibility of a $W_B \rightarrow W_A^{\bullet +}$ electron transfer (transfer II in Fig. 4.1) succeeding the $W_A \rightarrow FAD^*$ electron transfer (transfer I in Fig. 4.1). The rate of transfer II was shown to be crucial as this transfer has to outrun flavin protonation. Once flavin protonation takes place the closely spaced [$FADH^{\bullet}$ + $W_A^{\bullet}$] radical pair becomes stabilized and impedes transfer II energetically.

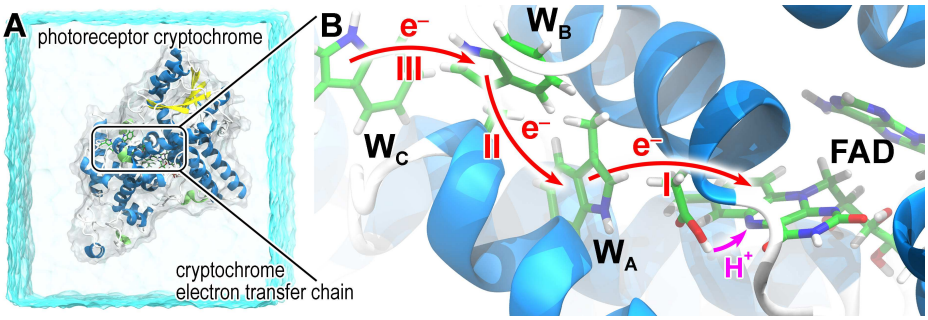


Figure 4.1: Cryptochrome structure and active site. ***A.*** *Structure of the protein cryptochrome-1 from Arabidopsis thaliana [145] in water environment.* ***B.*** *Cryptochrome internally binds the flavin cofactor, which governs its signaling through light-induced electron (red arrows) and proton (magenta arrow) transfers involving a chain of three tryptophan residues, W_A, W_B, W_C. The three consecutive electron transfers $W_A \rightarrow FAD^*$, $W_B \rightarrow W_A^{\bullet+}$, and $W_C \rightarrow W_B^{\bullet+}$ are labeled I, II, and III, respectively. Adapted from [146].*

In this chapter the electron transfer along the tryptophan triad in *A. thal.* cryptochrome is investigated by means of the direct electron transfer approach, which allows to take into account the complete biological environment of the active site. These environment components have never been accounted for in earlier computations. Unbiased dynamic simulations of the electron transfer along the tryptophan triad are performed in order to (i) identify the tryptophan which together with the FAD forms a persistent radical pair, and (ii) to reveal the underlying molecular mechanism of the photo-activation reaction in plant cryptochrome.

Similarly to the description of the electron transfer process, a completed inclusion of the protein environment and solvent is necessary in order to obtain a balanced description of the proton transfer energetics and to unambiguously determine whether the electron transfer along the tryptophan triad exceeds flavin protonation. Free energy calculations are used to study the protonation reaction of flavin in greater detail, with account for its complete biological environment. Ultimately, the results will be discussed in terms of the magnetoreception hypothesis.

Parts of this chapter are published as "'Solvent Driving Force Ensures Fast Formation of a Persistent and Well-Separated Radical Pair in Plant Cryptochrome"' in the Journal of the American Chemical Society [146].

4.1 Simulation Setup

4.1.1 MD Simulation of the Cryptochrome Resting State

The protein structure of *A. thal.* cryptochrome-1 used in this study is based on the X-ray crystal structure by Bräutigam *et al.* (PDB ID 1U3C) [145]. The FAD cofactor was built from riboflavin

and adenosine diphosphate (ADP) models using a force field parametrization developed earlier [130, 132]. Atomic charges for both neutral FAD and negatively charged $FAD^{\bullet-}$ were determined from RESP calculations [133] and are provided in the appendix E. The *in situ* pKa value of the D396 side chain, determined with a MCCE [148] calculation was about 10. Therefore, D396 was modeled as protonated. The protein model was solvated in a rectangular periodic box of 93 × 101 × 90 $Å^3$ size filled with TIP3P water molecules [149], neutralized with Na^+ ions. Excess Na^+ and Cl^- were added to obtain a 50 mM solution of NaCl. All simulations were performed with the GROMACS package [150]. The improved Amber99 force field was used [130, 131] for the protein parts. The structure of cryptochrome in its resting state obtained after a 100 ns MD simulation performed in an earlier investigation [67] was used as a starting configuration for the present simulation. These 100 ns were thus considered as an extended equilibration of the system, and were performed using the NAMD package [151] with the CHARMM27 [152, 153] forcefield. Finally, a 170 ns MD simulation was carried out. The integration time step was 2 fs. The temperature was kept at 300 K by applying the Nose-Hoover thermostat [101] while the pressure was controlled by the Parinello-Rahman Barostat [102] at a value of 1 atm. The LINCS algorithm [134] was employed to keep bonds involving hydrogen atoms at a fixed length. The time evolution of the cryptochrome backbone root mean square deviation (RMSD) is shown in the appendix A, confirming the stability of the simulation.

4.1.2 Direct Electron Transfer Simulation

QM/MM electron transfer simulations and QM/MM umbrella sampling simulations were performed with an in-house version of GROMACS 4.6 for the MM part and DFTB [93, 154] for the QM part. Force field parameters for the tryptophan cation radical were obtained as described before [126]. A time step of 1 fs was used in all QM/MM simulations. The temperature was kept at 300 K by applying the Nose-Hoover thermostat [101] while the pressure was controlled by the Parinello-Rahman Barostat [102] at a value of 1 atm.

In the electron transfer simulations the QM region includes the amino acid side chains in the *A. thal.* cryptochrome that are involved in the electron transfer directly, i.e. the side chains of tryptophans W400, W377 and W324.. For the sake of simplicity, in the following are denoted as W_A, W_B, and W_C, respectively, as also noted in Fig. 4.1B. The corresponding radical pairs [$FAD^{\bullet-}$+W400$(H)^{\bullet+}$], [$FAD^{\bullet-}$+W377$(H)^{\bullet+}$] and [$FAD^{\bullet-}$+W324$(H)^{\bullet+}$] are thus labeled RP-A, RP-B and RP-C, respectively. Snapshots from the MD simulation of the cryptochrome in its resting state served as starting structures for the electron transfer simulations. In total 32 snapshots were chosen from the time interval between 125 ns and 165 ns.

4.1.3 Potential of Mean Force Calculations

The free energy of FAD protonation via D396 was obtained with umbrella sampling simulations followed by a weighted histogram analysis (WHAM) [103]. The QM region includes the isoalloxazine ring of the FAD and the side chain of ASP396. The starting geometries for the simulations in the individual umbrella sampling windows were taken from a QM/MM MD simulation performed with an additional harmonic potential that was being shifted gradually with a rate of 0.024 Å/ps. This slow pulling simulation gave rise to structures spanning the entire proton transfer path, avoiding unphysical distortions of the holo-protein structure. The reaction coordinate was defined as the distance between the hydrogen of D396 and N5 of the flavin group . The starting point of the free energy calculation corresponds to a conformation with the COOH group of D396 turned entirely towards the flavin co-factor. The reaction coordinate was sampled from 2.6 Å down to 0.9 Å to span the transfer of the proton towards the N5 atom. Eight umbrella sampling windows with a spacing of 0.2 Å between 2.6 and 1.7 Å, and, additionally, seven umbrella sampling windows with 0.1 Å spacing from 1.6 to 0.9 Å were used. The force constant of the biasing potential was set to 100,000 kcal/mol/nm^2, and a QM/MM MD simulation of 150 ps was performed for each window.

4.2 Electron Transfer Dynamics

4.2.1 Analysis of Individual Electron Transfer Simulations

The ensemble of structural cryptochrome models was used as a starting point for 32 electron dynamics simulations. For each electron dynamics simulation the time evolution of the radical pair population was recorded. The population of a radical pair $Q = |a_m|^2$ (Eq. 2.62) defines which tryptophan is forming a radical at a given time instance. Per definition, radical pair population varies between 0 and 1, where 0 is characteristic for a neutral tryptophan, while 1 denotes the cation radical state. Initially, cryptochrome is assumed to be in the RP-A state, as the RP-A state is formed in less than a picosecond after flavin photo-excitation [66]. In the course of the simulation, the FAD cofactor remains negatively charged while the electron is free to move between W_A, W_B and W_C.

The analysis of the individual trajectories showed that the electron transfers $W_B \rightarrow W_A(H)^{\bullet+}$ and $W_C \rightarrow W_B(H)^{\bullet+}$ are completed in less than 50 ps for the majority of trajectories. For 6 trajectories out of 32, transfer from W_C took up to 100 ps. For half of the trajectories, occasional, short reverse transfers to W_B or W_A were observed. Transfer of the electron from W_B and no further transfer from W_C was found for 2 out of 32 trajectories. Examples of fast and comparably slow transfer as well as an example without transfer to W_C are shown in Fig. 4.2.

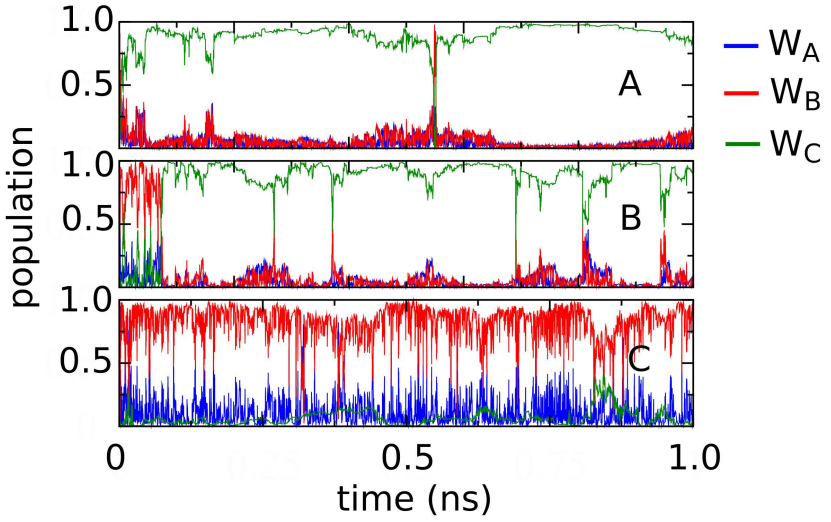


Figure 4.2: Characteristic examples of individual electron transfer trajectories. *The three panels show the time evolution of the population of each of the three tryptophans W_A, W_B, W_C over an interval of 1 ns for fast (**A**), slow (**B**) transfers. **C** shows the evolution of the population only up to $W_B(H)^{\bullet+}$. For **C** the population of both W_A and W_B frequently amounts to values of 0.5. This can be interpreted as frequent delocalization of the electron between W_A and W_B.*

4.2.2 Radical Pair Population Obtained from Ensemble Average

To account for the variability observed in the individual simulations and to obtain a clearer physical picture of electron transfer reactions we averaged over the ensemble of simulations.

Figure 4.3 summarizes the results of electron transfer dynamics in cryptochrome and shows the average population of the radical pair states RP-A, RP-B and RP-C over an interval of 1 ns. As follows from Fig. 4.3, the radical pair RP-A (blue) decays quickly, within 20 ps, giving rise to the formation of the RP-B state (red) and finally the RP-C state (green). The radical pair RP-B is only built up to about 40% when it starts decaying, giving rise to the population of the RP-C state. After 150 ps, the RP-C population remains stable at about 80% for the rest of the simulation. The observed behavior can be understood as a two-step electron transfer process with the transfer $W_B \rightarrow W_A(H)^{\bullet+}$ followed by a second transfer $W_C \rightarrow W_B(H)^{\bullet+}$ yielding RP-C as the ultimate radical pair.

After 1 ns, about 10% of RP-B still remains, indicating that there is a small but significant probability of RP-B stabilization. This observation is also supported by an analysis of the individual simulations, which showed that for 2 out of 32 simulations the electron has transferred from W_B to $W_A^{\bullet+}$ but not transferred on from W_C to $W_B^{\bullet+}$;

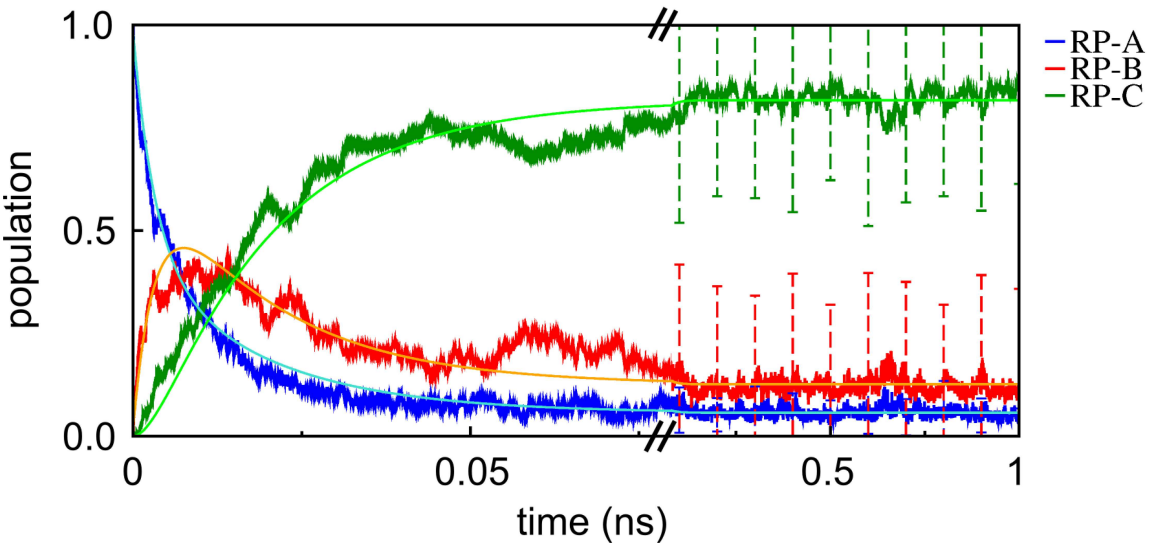


Figure 4.3: Formation and decay of three possible radical pair states in A. thal. cryptochrome. *The initially occupied RP-A state (blue) decays rapidly giving rise to the RP-B state (red) followed by the formation of the RP-C state (green). The radical pair population is obtained as an ensemble average of 32 simulations of 1 ns length which were calculated from snapshots taken from a MD simulation of the cryptochrome resting state at 500 ps intervals to represent the ensemble of cryptochrome structural variety. Error bars denote the standard deviation and indicate the variance of the time evolution of the population observed for the underlying individual simulations. The average populations (light blue, orange, light green lines) represent the fitting curves from a two-step kinetic model, Eq. (2.64-2.66), which has been fitted to the simulation data.*

The results from the electron transfer dynamics simulation are in an excellent agreement with mutational studies of *A. thal.* cryptochrome-1 [144], which show a prominent role of the tryptophan triad for electron transfer and identify W_C as a likely candidate for the terminal electron donor. A residual activity of the W324 mutant observed in the experiment might be explained by the small but existing RP-B population seen in the present simulations.

4.2.3 Radical pair separation

The spatial separation of the radical pair is a crucial parameter for the weak magnetic field sensitivity [155], as the dipole-dipole and exchange interactions between the radical partners become less relevant with increasing distance. Table 4.1 summarizes the distances between the flavin molecule and the tryptophanes W_A, W_B and W_C for different redox states of the protein. The distances obtained for the ground state of the protein are in agreement with the values calculated

from the crystal structure [156].

Table 4.1: Center to center distances between the flavin isoalloxazine ring and the tryptophan side chains W_A, W_B and W_C. *Calculated for different redox states of the protein. The resting and the radical pair states RP-A, RP-B and RP-C are considered. Bold values indicate the distances between $FAD^{\bullet-}$ and $W^{\bullet+}$ radicals. The distances are provided in Å.*

	d(FAD...+W_A)	d(FAD...+W_B)	d(FAD...+W_C)
resting state	8.2± 0.2	13.5 ± 0.3	18.7 ± 0.3
RP-A state	**7.9± 0.3**	13.5 ± 0.4	18.5 ± 0.4
RP-B state	9.3± 0.2	**12.2 ± 0.4**	16.7 ± 0.3
RP-C state	8.3± 0.2	13.6 ± 0.3	**19.4 ± 0.4**

The radical separation for RP-B is 12.2 Å while for RP-C it is 19.4 Å on average. The distance d(FAD...+W_B) depends on the redox state of the protein and for the RP-B state turns out to be about 1.3 Å lower as compared to the resting state of cryptochrome, while the distance d(FAD...+W_C) for the RP-C state of cryptochrome increases by about 0.7 Å. The changes in the radical pair distances can most likely be attributed to a small rearrangement of the tryptophanes due to the solvent polarization by the moving electron. Most importantly, both RP-B and RP-C exhibit a radical pair separation above the critical distance [72, 147], with RP-C being much more preferable for magnetoreception due to the larger distance.

4.2.4 Electron Transfer Rates

To connect the results from the performed simulations with the available experimental data, a numerical fit (see light color lines in Fig. 4.3) of the average radical pair population according to the kinetic model outlined in Sec. 2.3.3.4 was conducted

$$\mathrm{W_A} \underset{k_b^{II}}{\overset{k_f^{II}}{\leftrightarrows}} \mathrm{W_B} \underset{k_b^{III}}{\overset{k_f^{III}}{\leftrightarrows}} \mathrm{W_C} \tag{4.1}$$

where k^{II} and k^{III} denote the rate constants for electron transfer steps II and III, respectively, see Fig. 4.1B, and the subscripts f and b indicate the forward and backward transfers, respectively. The final numerical fit of the radical pair population is shown in Fig. 4.3 with light color lines while the obtained rate constants are summarized in Tab. 4.2.

The rates of the electron transfers $\mathrm{W_B} \rightarrow \mathrm{W_A^{\bullet+}}$ and $\mathrm{W_C} \rightarrow \mathrm{W_B^{\bullet+}}$ are found to be 190 $\mathrm{ns^{-1}}$ and 70 $\mathrm{ns^{-1}}$, respectively. The rate $k_{W_B \rightarrow W_A^{\bullet+}}$ is more than twice as large as the rate $k_{W_B \rightarrow W_C^{\bullet+}}$. This indicates a two-step electron migration with a fast transfer $\mathrm{W_B} \rightarrow \mathrm{W_A^{\bullet+}}$ followed by a somewhat slower, but still fast transfer $\mathrm{W_C} \rightarrow \mathrm{W_B^{\bullet+}}$.

The mean-field approach, used in our model to propagate the electron, suffers from electron

Table 4.2: Rate constants of the electron transfer steps in A. thal. cryptochrome. *The error bars denote the deviation of the numerical fit from the simulation data.*

e^- transfer step	donor acceptor	k_f [ns^{-1}]	k_b [ns^{-1}]
II:	$W_B \rightleftarrows W_A$	190± 18	85± 18
III:	$W_C \rightleftarrows W_B$	70± 18	10± 18

over-delocalization, an effect that gradually increases over the course of a simulation. As a result the population of the less probable states RP-A and RP-B will be overestimated while the population of the more probable state RP-C will be underestimated at the end of the 1 ns simulations. The over-delocalization, thus, results in an underestimated average population of RP-C and could lead to an overestimate of the electron transfer rates by about one or two orders magnitude of magnitude.

However, the derived electron transfer rates seem to be in an excellent agreement with the experimental kinetics obtained by Immeln *et al.* [66], suggesting electron transfer from a tryptophan residue on the protein surface to the flavin to be completed by a 100 ps delay. The experimental kinetics [66] yielded three time constants. The time constant of 0.4 ps was assigned to the initial tryptophan-to-flavin electron transfer (transfer I in Fig. 4.1B). The other two time constants 4 – 15 ps (rate constant 67 – 250 ns^{-1}) and 30 – 50 ps (rate constant 20 – 33 ns^{-1}) were both related to the rate of appearance of the tryptophan cation radical species. Comparison to the electron transfer rates from the kinetic model indicates that these two experimental rate constants could be related to the faster electron transfer $W_B \rightarrow W_A^{\bullet+}$ followed by the comparably slower electron transfer $W_C \rightarrow W_B^{\bullet+}$ along the tryptophan triad.

Maeda *et al.* [68] obtained an estimate for the backtransfer rate k_b^{III} for transfer III (0.01 ns^{-1}) from fitting the experimental data to a quantum spin dynamics model, assuming that k_b^{III} is close to the singlet triplet dephasing rate. This estimate suggests that the backtransfer rate k_b^{III} derived in the present study is overestimated, as it turns out to be three orders of magnitude higher. As argued above, the present calculation indeed could lead to a somewhat larger value of the backtransfer rate due to the electron over-delocalization arising in the mean-field approach. The estimated backtransfer rates should therefore be treated as a lower thresholds for the actual kinetic rates. However, the obtained values are still noticeably smaller than the forward rates, which suggests that electron backtransfer would be strongly suppressed in cryptochrome.

4.3 Potential Alternative Electron Donor

An alternative electron transfer donor was proposed based on the relative orientation of the flavin and tryptophan side chains. An estimate of their optimal angle for radical pair formation was deducted from the anisotropy values collected during femtosecond spectroscopy measurements [66].

Analysis of the *A. thal.* crystal structure indicated that W_{379}, which is not part of the tryptophan triad, is oriented at the most favorable angle and distance relative to the flavin. It was pointed out, that according to mutational studies of *A. thal.* cryptochrome that W_C is a likely candidate for the terminal tryptophan, but also that W_{379} has not been investigated yet [66, 144].

Inclusion of the W_{379} residue as electron transfer site during the direct electron transfer dynamics allows to test this hypothesis and to investigate the effect of W_{379} on the electron transfer dynamics. Fig. 4.4 shows the ionization potentials of the three tryptophans of the triad and the additional W_{379} residue in the $W_{379}^{\bullet+}$ radical state. W_{379} is stabilized by the environment by more than 1.5 eV compared to the other electron transfer sites. This indicates that W_{379}, in principle, can form a stable trap state.

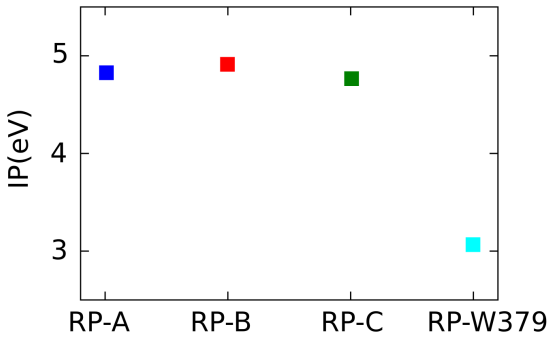


Figure 4.4: Ionization potentials of the electron transfer sites in the $W_{379}^{\bullet+}$ radical state. *The charge carrying site W_{379} residue has the lowest potential energy, indicating that it could in principle form a stable radical cation in less than 300 ps. The ionization potential was obtained from a classical MD simulation of the $W_{379}^{\bullet+}$ radical state.*

26 electron transfer simulations were carried out to probe the dynamic population of the [$FAD^{\bullet}$+$W_{379}^{\bullet+}$] radical pair. The results from the electron transfer dynamics simulations including W_{379} in addition to the main tryptophans W_A, W_B, W_C are shown in Fig. 4.5. W_{379} is never populated during the dynamics and the average population of the radical pairs RP-A, RP-B and RP-C seems unchanged compared to the simulations in which W_{379} was not considered (see Fig.4.3).

The time evolution of the electron transfer parameters, the ionization potential of each electron transfer site and the electronic coupling between two sites is summarized in Fig. 4.6. The ionization potentials of RP-B and RP-C decrease concurrently during the first 20 ps, yielding RP-C as the most strongly stabilized radical pair at the end of the simulation. The ionization potential of $W_{379}^{\bullet+}$ is never the lowest, $W_{379}^{\bullet+}$ remains energetically separated by more than 0.5 eV from RP-B and RP-C during all times. This indicates that electron transfer from W_{379} to $W_C^{\bullet+}$ or $W_B^{\bullet+}$ is highly unfavorable.

The average electronic coupling between W_A and W_B as well as the electronic coupling between

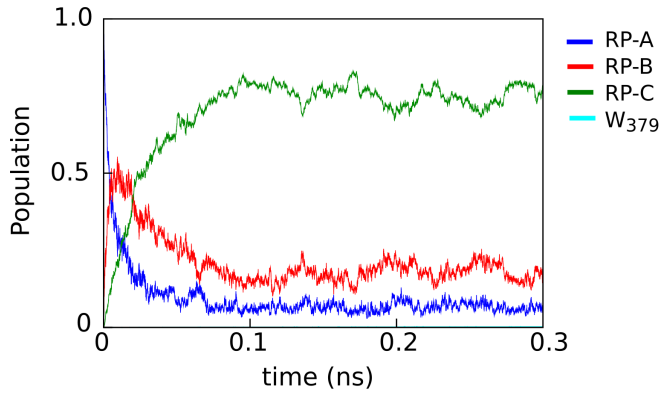


Figure 4.5: Population of the electron transfer sites. *W_{379} is never populated, while the population dynamics of the other radical pairs seems unchanged. RP-A decays quickly, giving rise to transient population of RP-B followed by formation of the RP-c radical pair. The population is obtained as an ensemble average over 26 simulations of 300 ps length, which were taken from an MD simulation of the cryptochrome resting state.*

Table 4.3: Pairwise distances between all potential electron transfer sites. *Obtained from 1 ns length MD simulations. All distances are given in Å.*

	W_B	W_C	W_{379}
W_A	7.87		
W_B		7.52	10.63
W_C			8.69

W_B and W_C are found to be about 8 meV and 2.5 meV, respectively. The pairwise electronic coupling between all other electron transfer sites is close to zero, indicating that the probability for electron transfer to $W_{379}^{\bullet+}$ from any other site is negligible.

The center to center distance between the side chain of W_{379} and the other tryptophans in the cryptochrome resting state is summarized in Tab. 4.3

The distance d(W_C...+W_{379}) is 8.7 Å on average and is slightly larger than the distances d(W_A...+W_B) and d(W_B...+W_C). A small but significant electronic coupling was found in *E. coli* for residues separated by about 9 Å. The negligible electronic coupling of W_{379} is thus surprising. Visual inspection of the protein environment of W_{379} and W_C using VMD [18], revealed that a proline (P376) residue is situated between the two tryptophan side chains, blocking the electron transfer. The relative orientation of the tryptophan residues and the proline residue is shown in Fig. 4.7. If proline acts as a bridge residue and mediates the tunneling electron transfer between W_{379} and W_C, cannot be probed by the direct electron transfer approach. Therefore, the involvement of W_{379} as terminal electron donor cannot be completely ruled out.

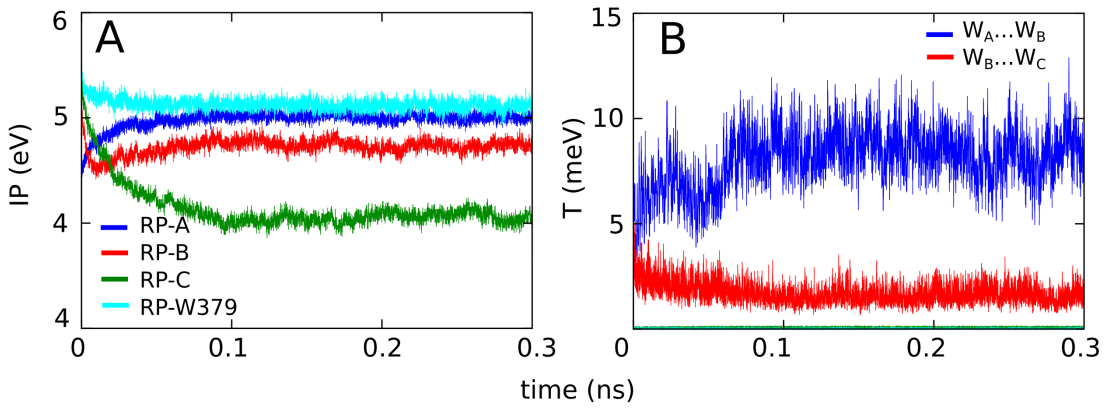


Figure 4.6: Evolution of the electron transfer parameters during the electron transfer dynamics. ***A*** *The ionization potential of the additional electron transfer site W_{379} is energetically separated from the ionization potential of the main sites.* ***B*** *The electronic couplings indicate a high probability for electron transfer only between the main tryptophans. The probability for electron transfer to $W_{379}^{\bullet+}$ from any other site is negligible.*

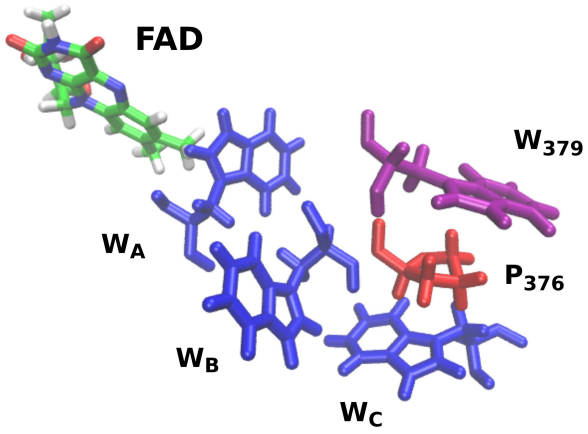


Figure 4.7: Cryptochrome active site including potential alternative terminal electron donor. *W_{379} (purple) has been proposed as alternative terminal electron donor. Electron transfer to the chain of the three conserved tryptophan (blue) via W_C is impaired by a proline residue (red). Protein and solvent environment are omitted.*

4.4 Driving Force for Electron Transfer Along the Tryptophan Triad

The cryptochrome resting state does not have an appropriate potential energy surface, brought about, e.g., by a specific arrangement of charged amino acids around the tryptophan triad, that would energetically favor the sequential electron transfer. This has already been demonstrated on a purely quantum mechanical model of the active site, where rearrangements induced by charge separation were shown to be of crucial importance [67, 147]. This indicates that there is no permanent, resting state electrostatic potential resulting from protein and solvent electrostatic interactions. The electron transfer driving force rather occurs due to dynamic polarization of the environment by the electron.

To quantify the stabilization of each radical pair by the environment, we computed the relative energy of each radical pair with respect to the energy of the FAD-tryptophan pair in the charge neutral state. Figure 4.8 depicts the relative energies averaged over 1 ns, including both protein and solvent environment (Fig. 4.8A), and considering the protein environment only (Fig. 4.8B). Figure 4.8A shows that all three radical pairs are stabilized by the polarized environment by more than 1.5 eV upon charge separation. However, the comparison of RP-A, RP-B and RP-C shows that the stabilization of each of the three radical pairs is different, with RP-C experiencing the strongest stabilization. The relative energy of RP-B is 0.44 eV lower compared to RP-A, and RP-C is even more stabilized by nearly 0.8 eV with respect to RP-A. These energy differences represent a driving force for electron transfer from W_C over W_B to $W_A^{\bullet+}$.

A similar effect has been reported for *E. coli* photolyase [126], and it was found to result from stronger solvation of the outer tryptophan which induced a stronger stabilization of the tryptophan cation radical due to a more favorable solvent repolarization.

To assess the effect of the solvent on electron transfer in cryptochrome in more detail, we excluded the QM/MM interaction of the water molecules and considered only the interaction with atoms from the protein environment, see Fig. 4.8B. A completely different picture emerges: RP-A is stabilized most strongly, while RP-B and RP-C experience less stabilization, suggesting that the electron transfer steps II and III in Fig. 4.1B would be hindered energetically. A similar observation was reported for cryptochrome using a quantum chemical approach for an active site model, i.e. neglecting the solvent and large parts of the environment [67]. In this case, an endothermic electron transfer process $W_B \rightarrow W_A^{\bullet+}$ was found, with a reaction energy of about +0.26 eV.

The performed analysis reveals clearly that the solvent plays an important role in *A. thal.* cryptochrome and is responsible for the strong driving force and exothermicity of the electron transfer process, inducing a predominant stabilization of the RP-C state.

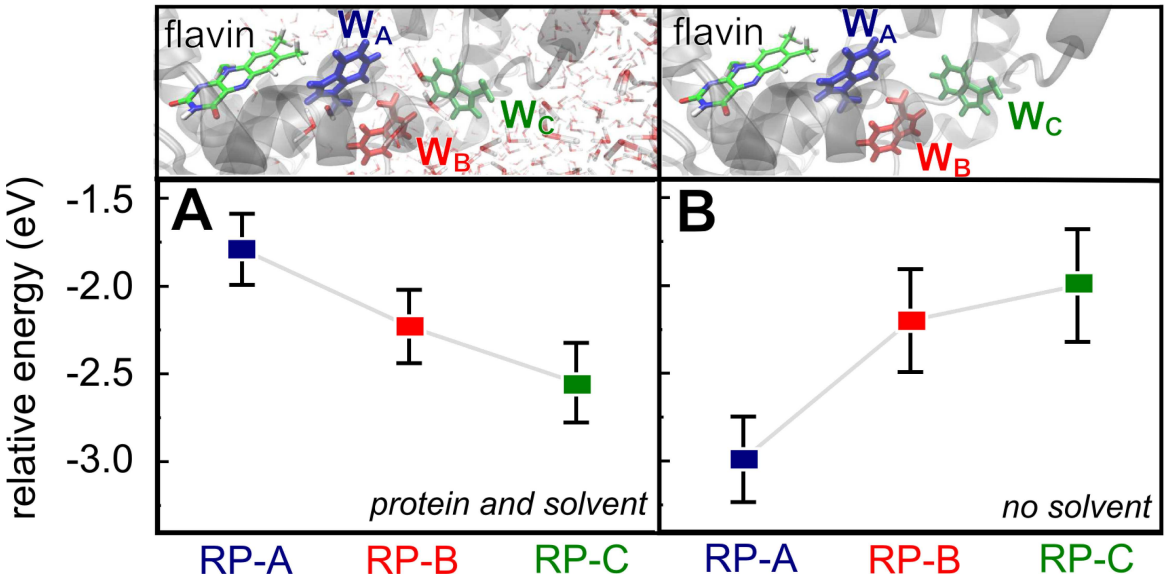


Figure 4.8: Contribution of the environment to the stabilization of each radical pair with respect to the charge neutral state. ***A:*** *The simulations including both protein and solvent environment show the average relative energy of the radical pairs to steadily decrease from RP-A to RP-C.* ***B:*** *The average relative energy due to protein environment only, neglecting the effect of the solvent, increases from RP-A to RP-C. Since the relative energy is strongly coupled to protein and solvent environment, large fluctuations occur. Error bars denote the standard deviation.*

4.5 Comparison to *E. coli* Photolyase

Experimental results show that *E. coli* photolyase and *A. thal.* cryptochrome are activated similarly [157]. However, the electron transfer dynamics simulation of *A. thal.* cryptochrome indicated a small but significant population of the RP-B state. In contrast, no probability for RP-B population has been found for *E. coli* photolyase [126, 158]. Subtle structural variations in the active site might induce unique electron transfer characteristics. In the following structural features of *E. coli* photolyase and *A. thal.* cryptochrome will be compared.

4.5.1 Solvent Exposure of Individual Tryptophans

Analysis of the relative energies of the radical pair states reveals that the photo-activation reaction in both proteins is driven by differential solvation of the thryptophans. Comparing the degree of solvent exposure of the three tryptophans for cryptochrome and photolyase, a small but crucial difference is revealed. Photolyase shows a larger difference in solvation of W_B and W_C than ob-

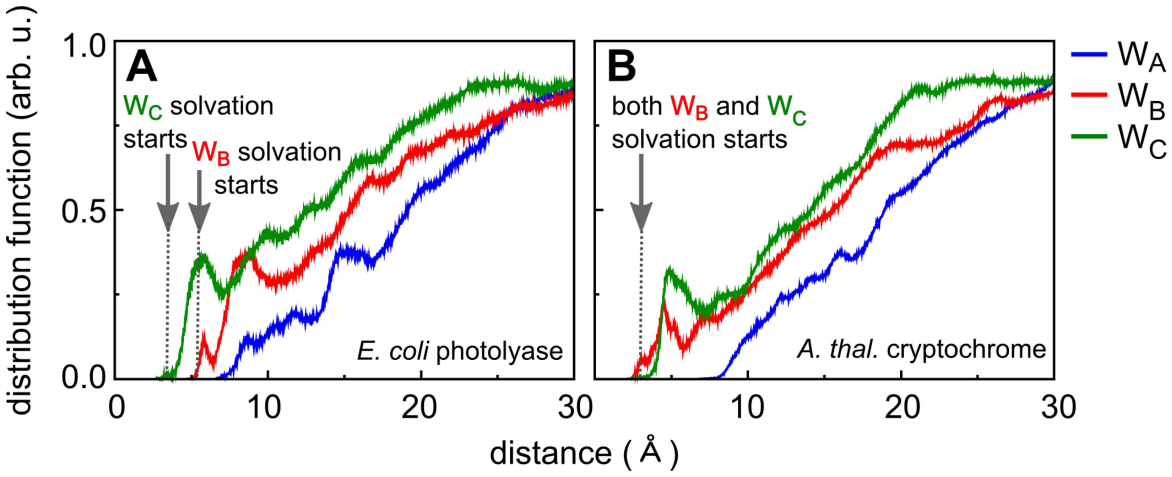


Figure 4.9: Radial distribution of water around the three tryptophans. *The radial distribution function is a measure for the solvent accessibility of each tryptophan W_A (blue), W_B (red) and W_C (green) in the resting state for E. coli photolyase (**A**) and A. thal. cryptochrome (**B**). W_B is found to be better solvated in the case of A. thal. cryptochrome than in the case of E. coli photolyase. The Gromacs tool g_rdf was used to calculate the radial distribution functions.*

served in cryptochrome, with W_B being more buried inside the protein matrix. To quantify the solvent exposure of the individual tryptophans, it is useful to compute the radial distribution of water molecules around the three tryptophans. The radial distribution function defines the probability of observing a water molecule within a sphere of radius r from the respective tryptophan W_A, W_B or W_C. It is obtained by counting the number of water molecules found within this sphere for each time step along a MD trajectory, and normalizing it by the number of water molecules found in a sphere with a sufficiently large cut-off radius (here, 30 Å is used).

Figure 4.9 shows the radial distribution of water molecules around the three tryptophans of photolyase [126] and the data calculated for cryptochrome. In the case of photolyase, the water distribution sets on at 5.5 Å for W_B and at 4 Å for W_C. In contrast, the distribution of water for cryptochrome tryptophans sets on at 4 Å for both residues. The higher maximum value for W_C in the case of cryptochrome indicates that a larger number of water molecules can be found at a distance up to 5 Å from W_C than in the vicinity of W_B. This still renders W_C the most solvated tryptophan of the triad in cryptochrome also, and explains qualitatively why RP-C is stabilized the best. However, the merely qualitative difference between RP-B and RP-C also hints to the non-zero probability of RP-B stabilization found in the dynamical electron transfer simulation for cryptochrome, which was never observed in *E. Coli* photolyase [126].

4.5.2 Structural Mechanism for RP-B Stabilization

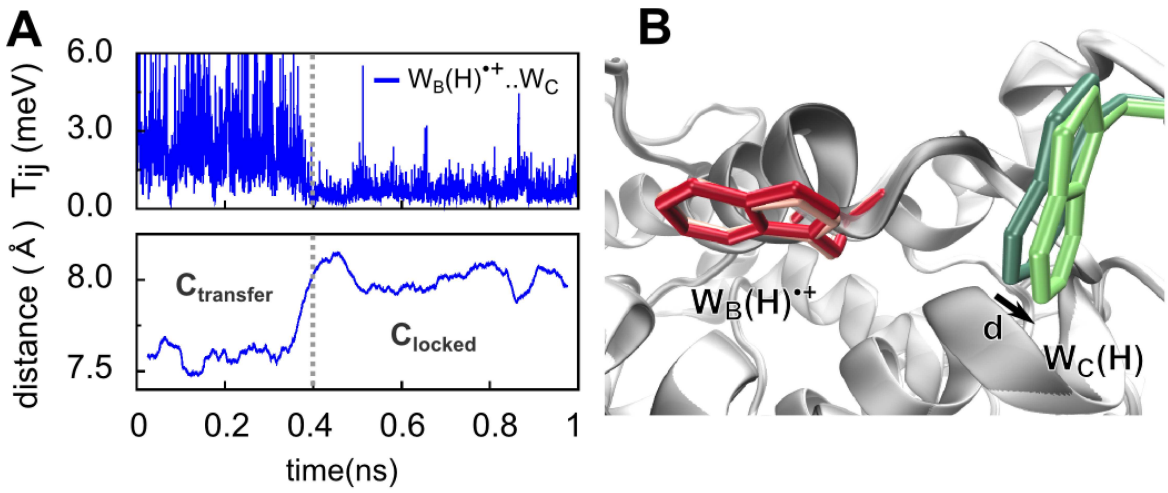


Figure 4.10: RP-B could be stabilized by a subtle rearrangement of W_C. A: *After about 400 ps, the distance between the center of mass of $W_B(H)^{\bullet+}$ and W_C increases (bottom), and the electronic coupling between $W_B(H)^{\bullet+}$ and W_C decreases to almost zero (top). The decrease of the electronic couplings implies a negligible probability for the electron transfer $W_B \leftarrow W_C(H)^{\bullet+}$.* ***B:*** *The increase of the distance between $W_B(H)^{\bullet+}$ and W_C gives rise to two conformations $C_{transfer}$ (light colors) and C_{locked} (dark colors). The two conformations $C_{transfer}$ and C_{locked} are obtained as average of the trajectory before and after 400 ps, respectively.*

For those simulations where RP-B became stabilized, and no $W_C \rightarrow W_B(H)^{\bullet+}$ electron transfer was possible, an increase of the center of mass distance between $W_B(H)^{\bullet+}$ and W_C by about 0.5 Å is observed, and the electronic coupling between $W_B(H)^{\bullet+}$ and W_C shows a step-like decrease to nearly zero as can be seen in Fig. 4.10A. The electronic coupling determines the probability of electron transfer between the tryptophans; a decrease of the electronic coupling means a drop of the probability of electron transfer between $W_B(H)^{\bullet+}$ and W_C. The difference between the conformations of the tryptophan pair is very subtle, see Fig. 4.10B. Still it may affect the electron transfer substantially: the low electronic coupling renders the electron transfer from W_C to $W_B(H)^{\bullet+}$ almost impossible and provides a structural mechanism for the stabilization of RP-B against the solvent driving force. Note that electron back transfer $W_B(H)^{\bullet+} \rightarrow W_A$ is still possible but expected to be slow, due to the energetically more favorable stabilization of RP-B compared to RP-A, see Fig. 4.8.

No additional mechanism for the stabilization of RP-A and RP-C is observed for *A. thal.* cryptochrome. In a previous study on *E. coli* photolyase, a structural stabilization mechanism was discovered for RP-C [126]. Several nanoseconds after the RP-C was formed, a spontaneous increase

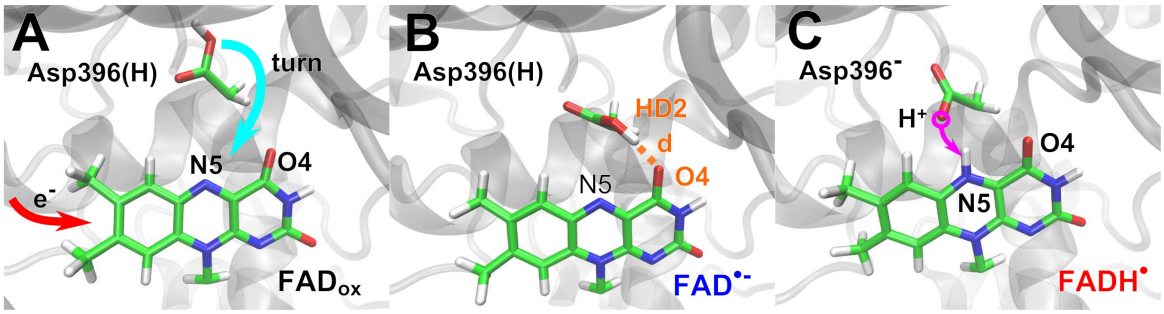


Figure 4.11: Protonation of flavin by D396. *Shown is the spontaneous rotation of the COOH group of D396(H) towards flavin subsequent to the radical pair formation that ultimately leads to flavin protonation. Relative orientations of D396 and flavin obtained through MD simulation of* ***A:*** *the resting state of cryptochrome and* ***B:*** *cryptochrome in the radical pair state [$FAD^{\bullet -}$+$W400(H)^{\bullet +}$]. A hydrogen bond (orange) is formed between $D396^{HD2}$ and $flavin^{O4}$. Protonated flavin* ***C*** *is obtained from a QM/MM umbrella sampling calculation.*

of the distance and decrease of the electronic coupling between W_B and $W_C(H)^{\bullet +}$ was observed, essentially blocking the possibility of the $W_C(H)^{\bullet +} \rightarrow W_B$ electron back transfer.

4.6 Protonation of Flavin via D396

Recently [67], it was shown that D396(H) is involved in a structural transformation once the flavin gains a negative charge and W_A becomes positively charged, and this might catalyze the protonation of flavin N5 via D396(H). In order to study the structural rearrangement subsequent to the formation of the primary radical pair [$FAD^{\bullet -}$+$W400(H)^{\bullet +}$], we conducted 14 classical MD simulations starting from different structures taken from the simulation of the cryptochrome resting state. The upper panel of Fig. 4.11 shows that after the formation of the radical pair, the COOH group of the D396(H) residue turns spontaneously towards the flavin, and the COOH group approaches the N5 and O4 atoms of the flavin group, see Fig. 4.11B. Thereafter, a hydrogen bond of a 1.8 Å length is formed between the O4 atom of the flavin and the COOH group of D396.

4.6.1 Benchmarking DFTB Proton Affinities

DFTB is able to compute relative proton affinities (PA) with sufficiently high accuracy compared to *ab initio* reference calculations, as described previously [93]. Since the FAD molecule was not included in the original test set [96], DFTB3 was tested for relative proton affinities of FAD and aspartate.

Vaccum proton affinities for the aspartic acid sidechain (ASP) as well as for the negatively charged flavin group $FAD^{\bullet -}$ are shown in Tab. 4.4. DFTB3 yields slightly lower values compared

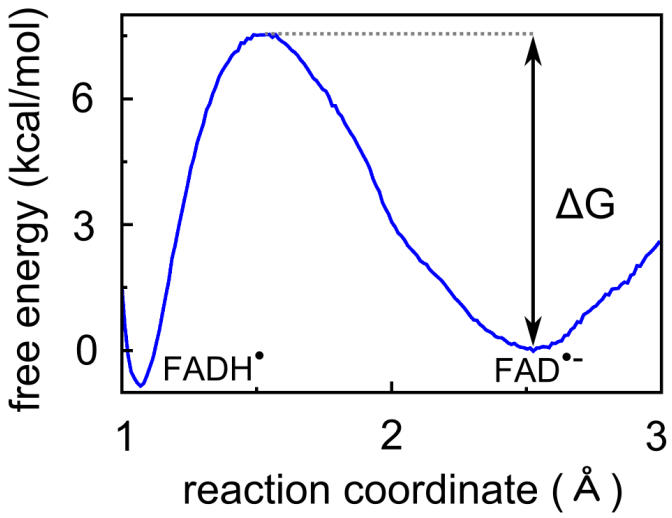


Figure 4.12: Free energy profile for protonation of flavin by D396. *Protonated flavin is obtained from a QM/MM umbrella sampling calculation, which yields the energy profile for flavin N5 protonation. The distance between the hydrogen of the COOH group and the flavin N5 atom was considered as the reaction coordinate.*

to DFT(PBE) and *ab initio* (MP2) results, especially for the flavin group. However, the relative proton affinities of the flavin group compared to the aspartic acid are comparable to those of the other two methods. Therefore, DFTB3 with the 3OB parameter set [96] is expected to accurately reproduce the reaction energy as well as the barrier heights for the reaction studied in this work.

Table 4.4: Vaccuum proton affinities. *Calculated with Gaussian03 [159] employing the aug-pvTZ basis set for wave function expansion. Energies are given in kcal/mol. Structure optimization was carried out with the B3LYP/TZVP method.*

	DFTB 3OB	PBE	MP2
ASP	-358	-352	-356
$FAD^{\bullet -}$	-341	-333	-335
rel. PA ASP ↔ FAD	-17	-19	-21

4.6.2 Potential of Mean Force Calculations

The free energy profile for the protonation reaction of the flavin by D396 was obtained from a QM/MM umbrella sampling simulation where the hydrogen of the COOH group of D396 was pulled towards the N5 of the flavin group. The distance between the hydrogen of D396 and N5 atom of the flavin group (see magenta arrow in Fig. 4.1B) was considered as the reaction coordinate and was sampled from 2.6 Å down to 0.9 Å to describe the proton transfer process completely.

The WHAM calculation yielded a PMF with a minimum for the D396 hydrogen bonded to the flavin O4 and another minimum, which was 1.5 kcal/mol deeper, located at the protonated flavin state. The free energy barrier for the proton transfer amounted to 7.5 kcal/mol, see Fig. 4.12.

4.6.3 Estimate of the Proton Transfer Rate

In the framework of the transition state theory [160], the rate for the proton transfer can be estimated as

$$k = \frac{k_B T}{h}\left(1 - e^{-\frac{h\nu}{k_B T}}\right) e^{-\frac{\Delta G}{k_B T}}, \tag{4.2}$$

where T denotes the temperature, ΔG is the free energy barrier, ν is the attempt frequency by which the transition state occurs, k_B is the Boltzmann constant and h is Planck's constant. The attempt frequency can be estimated as

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k_{H-N}}{m_H}}, \tag{4.3}$$

where $k_{H-N} \simeq 537 \frac{\mathrm{kcal}}{\mathrm{mol\AA}^2}$ is the H-N bond force constant (the characteristic value of the force constant in the CHARMM force field [152, 153]) , and m_H is the mass of the hydrogen atom. With Eq. (4.3) one thus obtains $\nu = 7.5 \times 10^{13}$ s^{-1}. Together with the free energy barrier of ΔG =7.5 kcal/mol, see Fig. 4.12, this leads to a rate constant of $k = 2.2 \times 10^7$ s^{-1}, which is equivalent to a time constant of 45 ns. This time constant can be understood as an estimate for the time it takes for the flavin to become protonated on average. In contrast, the electron transfer through the triad was found to proceed with a rate of about 2×10^{11} s^{-1} (a time constant of about 50 ps), as demonstrated above.

This analysis reveals that the proton transfer in *A. thal.* cryptochrome is expected to happen at least four orders of magnitude slower than the electron transfer. In a previous study [67], the protonation time was estimated to be about 1 ns, which is faster than the present estimate, but still significantly slower than the electron transfer discussed above. Femtosecond spectroscopy experiments observed the proton transfer to the flavin with a time constant of about 1.7 μs [64] and the electron transfer along the tryptophan triad to be completed within 100 ps [66]. While our theoretical investigation yields electron transfer rates in an excellent agreement with the experiment, the findings for the proton transfer rate should be considered qualitative. It should also be noted that the rate of flavin protonation apparently depends on various external factors, such as pH and the presence of charged co-factors around the cryptochrome active site [58, 59].

Note that the flavin protonation for the RP-A state was studied here in order to investigate if it could outrun the $W_B \rightarrow W_A(H)^{\bullet+}$ electron transfer process. However, since the electron transfer occurs significantly faster than the flavin protonation, the RP-C state will be formed, which would change the flavin protonation barrier. The barrier for flavin protonation in the RP-C state

will probably be higher than in the case of the RP-A state, as it will only be dominated by the pulling force from the negative flavin radical, while in the case of the RP-A state the positively charged $W_A(H)^{\bullet+}$ radical additionally accelerates the flavin protonation process. The performed analysis proves clearly that the electron and proton transfer processes in plant cryptochrome are separated in time by several orders of magnitude. The protein environment imposes an activation barrier on flavin protonation ensuring the formation of the much more persistent radical pair $FAD^{\bullet-}$+$W_C(H)^{\bullet+}$, which is well separated in space, before the flavin protonation could occur.

4.7 Conclusion and Discussion

Direct electron transfer simulations that include the complete protein and solvent environments, show that two sequential electron transfers $W_B \rightarrow W_A^{\bullet+}$ and $W_C \rightarrow W_B^{\bullet+}$ lead to the stabilization of the radical pair RP-C or RP-B within 150 ps. The computed electron transfer rates are in excellent agreement with the experimental kinetics obtained from femtosecond spectroscopy [66] and provide an in-depth understanding of the photo-activation mechanism of plant cryptochrome.

While the above analysis yields electron transfer rates that are in an excellent agreement with the experiment, the findings for the proton transfer rate can only be considered qualitative. It should also be noted that the rate of flavin protonation depends on various external factors, such as pH and the presence of charged co-factors around the cryptochrome active site [58, 59]. However, the potential of mean force calculations prove clearly that the electron and proton transfer processes in plant cryptochrome are separated in time by several orders of magnitude.

Analysis of the relative energies of the radical pair states RP-A, RP-B and RP-C reveals that the driving force results solely from differential solvation effects, i.e. from the fact that W_C is much more solvent exposed than W_A: the polarization of the solvent induces stabilization of RP-C by 0.8 eV with respect to RP-A. It is this strong exothermicity that renders the electron transfer to be very fast, stabilizes the radical pair and slows down electron back transfer reactions along the tryptophan triad.

A solvent-driven electron transfer along the tryptophan triad, inducing the stabilization of RP-C, was also found earlier in *E. coli* photolyase [126, 158]. Given the fact that the tryptophan triad is conserved throughout the cryptochrome/photolyase family, the findings suggest that the solvent driving force is a general principle governing the fast electron transfer process of the photoactivation reaction in members of the cryptochrome/photolyase family. It represents the main feature that allows the formation of a radical pair, comprising all the essential prerequisites (fast formation, well separated in space, stabilized to prevent back transfer) for the potential sensitivity to weak magnetic fields.

Furthermore, the present study establishes a functional role of solvent for biological reactions. For both *A. thal.* cryptochrome and *E. coli* photolyase we have found an additional structural

"locking" mechanism that enforces the electron reaction to be one-way, in addition to the stabilization already provided by the strong exothermicity of the reaction itself. This mechanism provides additional stabilization of RP-C on the nanosecond timescale in *E. coli* photolyase, while it allows stabilization of RP-B in *A. thal* cryptochrome, each time allowing to fine-tune the radical pair stabilization specific to the function of the individual protein.

No doubt that the proposed magnetoreceptive function of cryptochrome [8, 68, 161] is an intriguing property of the protein as it attracts the attention of an increasing number of researchers. It has been suggested [68] that an applied magnetic field influences the radical pair recombination reaction, which in turn competes with the tryptophan radical deprotonation reaction, see Fig. 1.4. Interestingly, the magnetic field effect in *A. thal.* cryptochrome was observed to be about three times larger compared to *E. coli* photolyase [68]. The present investigation provides a qualitative explanation to this observation.

The "locking" mechanism in *E. coli* photolyase leads to an additional, nanosecond stabilization of the RP-C state, which does not permit the electron back transfer easily, thereby rendering the influence of an applied magnetic field on the probability of radical pair recombination less important, as the radical pair recombination is significantly decreased. In contrast, the "locking" mechanism in *A. thal.* cryptochrome helps to stabilize RP-B, but it does not prevent the recombination of RP-C specifically. Thus, since radical pair recombination is possible in *A. thal.* cryptochrome, the magnetic field could induce an increased recombination of the radical pair, resulting in a higher sensitivity of *A. thal.* cryptochrome compared to *E. coli* photolyase. This is in accordance with the data of Maeda *et al.* [68] who estimated the rate constant for radical pair recombination to be 4.9×10^5 s^{-1} for *A. thal.* cryptochrome and 1.2×10^5 s^{-1} for *E. coli* photolyase. One might also speculate that the "locking" mechanism in *A. thal* cryptochrome is tailor-made to allow stabilization of RP-B without additionally preventing radical pair recombination. This potentially permits a comparably high sensitivity of *A. thal.* cryptochrome to weak magnetic fields.

CHAPTER 5

Re-evaluation of Electron Transfer Parameters of Aromatic Side Chains

Cryptochromes are thought to be activated by electron transfer from the protein surface to the flavin cofactor [66]. Experiments involving mutations of the tryptophan triad resulted in decreased flavin reduction in vitro but did not seem to eliminate biological activity of cryptochrome in vivo [56, 57, 162]. The discrepancy is suggested to be explained by the in vivo interaction of cryptochrome with small molecule activators such as ATP [58–60] which are supposed to facilitate the use of alternative electron transfer pathways [60] in addition to the tryptophan triad. Evidence for alternative electron transfer pathways involving tyrosine residues in *X. laevis* cryptochrome has been accumulated by time-resolved EPR spectroscopy experiments [163]. An alternate electron transfer pathway via the adenine moiety of FAD to a surface exposed tyrosine residue is implicated by mutational studies and EPR measurements of ATP-sensitive flavin reduction in cry2 from *A.thaliana.*

The significance of alternative electron transfer pathways in cryptochromes fueled the idea of refining the the DFTB-FO parameterization for tryptophan and tyrosine residues.

In order to describe the electron transfer energetics correctly, it is crucial to obtain the correct IP difference for the residues. A previous benchmarking study showed that DFTB reproduces experimental IP values for nucleobases with a deviation of 0.26 eV. IP differences between DNA bases were reproduced with a deviation of 0.23 eV, compared to 0.09 eV for DFT [123]. The IP difference for tryptophan and tyrosine sidechains was reproduced by DFTB with a similar accuracy [158].

The IP difference between neighboring side chains in photolyase have been found to be about 0.6 eV on average. Underestimation of the IP difference by about 0.2 to 0.3 eV might therefore

affect the electron transfer dynamics. This deviation can be corrected by application of an additive constant (shift) to the DFTB HOMO energies in such a way that the correct IP difference is obtained. The procedure to calculate the shift is explained and validated in the following.

5.1 Simulation Setup

Quantum chemistry calculations to obtain ionization potentials of tryptophan and tyrosine were carried out with Gaussian 03[159]. Only the side chains were considered in the quantum chemical description, with the C_β-atoms being terminated by hydrogen. Geometries were optimized at the B3LYP/6-31G** level. The single point energies were computed with the 6-31G** basis set for HF, B3LYP and PBE. Single point calculations with DFTB were carried out with the dftb+ program package [164] using the 3OB parametrization [93].

The effect of protein and solvent electrostatic potential on the ionization potentials was considered by including a distribution of point charges in the quantum chemistry calculations. Tryptophan and tyrosine geometries and the set of point charges were taken from an MD simulation of the cryptochrome resting state [165]. In order to account for the structural fluctuations of the system, single point calculations were carried out for 1000 snapshots of the MD simulation at the B3LYP/6-31G** level of theory using Gaussian03. DFTB single point energies in the presence of point charges were computed using the QM/MM interface of an in-house version of GROMACS 4.6 for the MM part and DFTB2 [154] for the QM part.

5.2 Test of the DFTB-FO Parametrization

The energetics of electron transfer between aromatic sidechains can be estimated from their IP difference. The first step is to benchmark DFTB against DFT and ab initio methods. The vertical IP can be obtained using the ΔSCF procedure which involves the calculation of the total energy of both the neutral $E^N$and the charged molecule E^{N-1}. A more efficient way to estimate the IP of a sidechain is the calculation of the HOMO energy ϵ_H. The HOMO energy is directly connected to the IP of a molecule via Koopmans' theorem (Eq. 5.1)for HF theory.

$$IP = E^{N-1} - E^N = -\epsilon_H \tag{5.1}$$

An equivalent theorem exists for DFT and is referred to as Janak's Theorem. It is important to note that Janak's theorem is true only if the correct exchange correlation functional was known. As the exchange-correlation functional is always approximated, the HOMO energies obtained with DFT are expected to deviate from the actual IP value. In practice, comparison of DFT HOMO energies with the experimental IP value for a large number of molecules yields an deviation of

about 2 eV [91].

Table 5.1 compares the IP values calculated with DFTB to those obtained with B3LYP, PBE and HF using the ΔSCF procedure. The vertical IPs calculated with DFTB match quite well with the experimental value. Both DFT functionals slightly underestimate the vertical IP compared to the experimental value. Values obtained with HF deviate from the experimental IP by more than 1 eV. The relative IPs which represent the difference between the IP of tryptophan and the IP of tyrosine are crucial for the electron transfer dynamics. The relative IPs obtained from the ΔSCF calculations agree well with experimental values.

The HOMO energies calculated with HF agree well with the experimental values while the HOMO energies obtained with DFT methods underestimate the experimental IP value significantly. However, the results are expected due to the approximation of the exchange correlation functional used in DFT. Most importantly, the relative HOMO energies obtained with the DFT methods agree much better with the relative IP obtained from experiment than the absolute values. Still, the relative HOMO energies are underestimated by about 0.2 to 0.3 eV compared to the relative IPs from the ΔSCF procedure obtained with higher-level methods and the experiment.

This deviation can be easily corrected by shifting the HOMO energy in such a way that the correct relative IP will be obtained. Practically, in an electron transfer simulation this is achieved by adding a constant value to the HOMO energy of one of the side chains.

Table 5.1: IP of tryptophan and tyrosine Sichechains. *Calculated as ΔSCF and ϵ_{HOMO} using DFT and HF methods (6-311G** basis set)*

	TYR		TRP		relative IP	
method	ϵ_{HOMO}	ΔSCF-v	ϵ_{HOMO}	ΔSCF-v	ϵ_{HOMO}	ΔSCF-v
DFTB	5.76	8.66	5.32	7.97	0.44	0.67
B3LYP	6.04	8.09	5.51	7.39	0.53	0.7
PBE	5.21	7.96	4.78	7.28	0.43	0.68
HF	8.28	-	7.59	-	0.69	-
exp,vert.		8.5		7.9		0.6
exp,adiabt.		8.0		7.2		0.8

5.3 Effect of Structural Variation on the Relative IP

As shown above the IP values calculated with DFTB agree well with those of higher-level methods. However, the gas phase calculations show that the relative IP calculated as HOMO energy of DFTB is slightly underestimated compared to IP values obtained by means of the ΔSCF procedure using higher level methods. This indicates that a shift in the range of 0.1 to 0.3 eV needs to be applied to the DFTB HOMO energy of one of the side chains to reproduce the electron transfer energetics energetics correctly.

As the electron transfer parameters have been shown to depend sensitively on the structural fluctuations [123, 166, 167], the IP values need to be evaluated along molecular dynamics trajectories. The effect of protein and solvent electrostatic potential on the ionization potentials is considered by including a distribution of point charges in the quantum chemistry calculations. Tryptophan and tyrosine geometries and the set of point charges were taken from an MD simulation of the cryptochrome resting state. The IP in the presence of point charges was calculated for 100 snapshots taken at 1 fs intervals. Fig. 5.1 shows the time evolution of the IP values computed by means of the ΔSCF procedure using B3LYP and obtained as HOMO energy using DFTB.

As discussed above, the HOMO energies represent a poor approximation of the absolute IP, deviating by about 2 eV from the IP value obtained with the ΔSCF procedure using B3LYP. For the tyrosine side chain the deviation is found to be slightly larger than for tryptophan side chain. However, the HOMO energies show identical time evolution and follow the fluctuations of the real IPs quite accurately. Since the relative fluctuations are crucial for the description of electron transfer, the dynamic features of the DFTB HOMO agree nicely with those of the DFT ΔSCF energy and are well suited to describe the energetics.

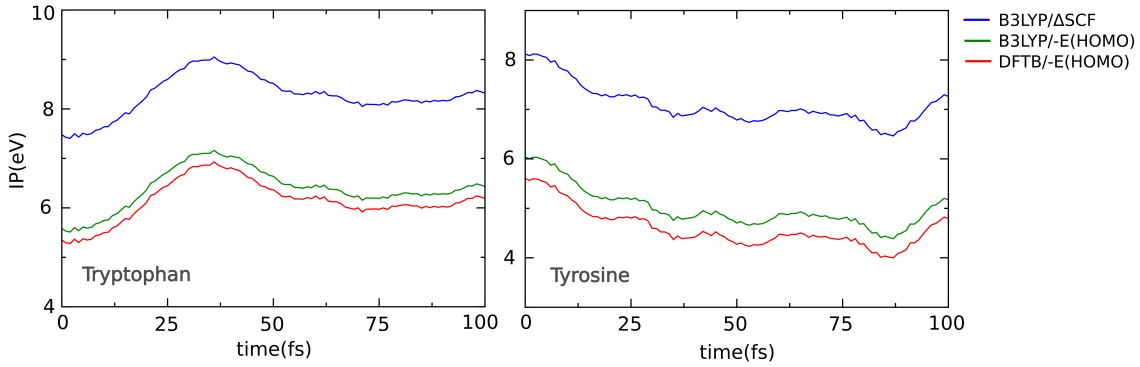


Figure 5.1: Comparison of the ionization potential calculated as ΔSCF and DFTB HOMO energy. *Obtained from a 100 fs molecular dynamics trajectory. The HOMO energies of tryptophan (left) and tyrosine (right) side chains follow the time evolution of the ΔSCF IP values quite accurately and are well suited to describe the relative energetics.*

In order to take into account the complete ensemble of cryptochrome structural variety, the quantum chemical calculations were carried out for 1000 structures and charges taken from snapshots of the MD simulation at an interval of 1 ps. Fig. 5.2 shows the relative IP of the tyrosine and tryptophan side chains obtained from the DFTB HOMO energies and the relative IP obtained by the means of the ΔSCF procedure. The relative IP values are found to depend strongly on the underlying structure from the molecular dynamics simulation. The relative IP values fluctuate with a standard deviation of 1.47 eV. As the dynamical features of the ΔSCF IP were reproduced

correctly by the DFTB HOMO energies as shown above, the relative IP obtained from the DFTB HOMO energies can be safely corrected by applying a shift to the HOMO energy of one of the side chains.

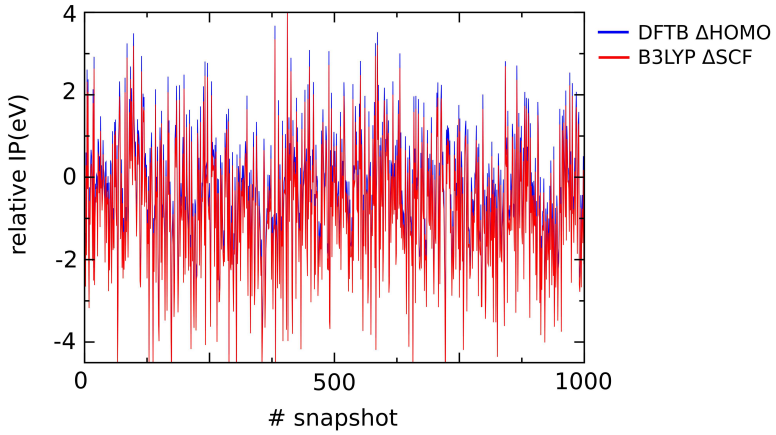


Figure 5.2: Relative ionization potential. *Calculated as ΔSCF and DFTB HOMO energy for 1000 snapshots of a molecular dynamics trajectory. The relative IP computed from the DFTB HOMO energies is slightly overestimated by about 0.3 eV compared to the ΔSCF relative IP. The correct relative IP can be obtained by applying a shift to the DFTB HOMO energy of one of the side chains.*

The procedure applied to obtain the energy correction is explained in the following: The IP value of a side chain obtained with the DFTB HOMO energy is assumed to deviate from the IP value calculated with the ΔSCF procedure by a constant value C. The constant C might be obtained by minimizing the deviation function of the DFTB HOMO energy and the ΔSCF IP

$$RMSD = \sqrt{\frac{\sum_i^N (IP_{HOMO} - C - IP_{\Delta SCF})^2}{N}}. \tag{5.2}$$

For the tyrosine side chain the numerical fit yields a minimal RMSD of 0.028 eV for C=-2.425 eV. For the tryptophan side chain the minimal RMSD of 0.014 eV is found for C=-2.136 eV. This means that by shifting the tyrosine DFTB HOMO energy by +0.289 eV the correct relative IP is obtained with DFTB. The minimal RMSD values indicate how well the structural fluctuations of the ΔSCF IP are reproduced by the DFTB HOMO energies.

5.4 Discussion and Conclusion

The gas-phase calculations of the ionization potential of tryptophan and tyrosin side chains show that the ionization potential obtained as HOMO energy is underestimated by several eV com-

pared to the experimental values. The relative IP, which is the important quantity for electron transfer, is only underestimated slightly by about 0.2 to 0.3 eV. This small deviation, however, might already affect the dynamics of electron transfer. The relative IP calculated by the means of the ΔSCF procedure agrees well with the experimental values and thus might be used as a reference to correct the DFTB HOMO energies.

As the dynamical features of the ΔSCF IP are well reproduced by the DTFB HOMO energies, the relative IP obtained with DFTB can be safely modified by applying a shift to the HOMO energies.

The shift is obtained by fitting the deviation of the DFTB HOMO energy to the ΔSCF energy obtained with DFT using the B3LYP functional for 1000 snapshots of a molecular dynamics trajectory. This procedure yields a shift of 0.289 eV for the tyrosine side chain. The minimum RMSD values indicate a negligible deviation of the structural fluctuations of the DFTB HOMO energy compared to the ΔSCF energy.

In conclusion, shifting the DFTB HOMO energy yields an improved description of the relative energetics of electron transfer in photolyases and cryptochromes. This makes the direct electron transfer approach a valuable tool in the search of alternate electron transfer pathways. The effect of the energy shift on the electron transfer dynamics and rates can be tested for example by applying the direct electron transfer approach to *X. laevis* cryptochrome.

CHAPTER 6

Development of Model Parameters for the Simulation of DNA Repair by *E. coli* Photolyase

The repair of radiation damaged DNA by the CPD photolyase [3] involves an electron transfer from the excited flavin cofactor to the thymine dimer lesion. This reaction creates a neutral semiquinone radical $FADH^{\bullet}$ and a radical anion thymine dimer $CPD^{\bullet -}$ which undergoes an ultra-fast disintegration. The repair mechanism has been studied extensively theoretically and experimentally by several groups, however, its exact nature remains elusive [23–27, 168, 169].

So far, the dynamic interaction of the electron with the protein and solvent environment has not been explicitly accounted for in the theoretical investigations. Solvation dynamics have been found to play a crucial role for the electron transfer along the tryptophan triad in photolyases and cryptochromes [126, 146, 158]. The direct electron transfer simulations allowed to study polaron formation induced by the moving electron and identified non-equilibrium effects to be important for the efficiency of the electron transfer reaction in *E. coli* photolyase. In both solved crystal structures of the photolyase-DNA complex, water molecules were found in the active site and the solvation of the thymine dimer in the binding pocket seems to be important for the stabilization of the thymine dimer radical anion after electron transfer from the $FADH^{-}$ [170]. The solvation dynamics at the active site have been suggested to modulate competing electron transfer channels such that the forward electron transfer from the excited $FADH^{-}$ is favored over electron transfer to the adenine moiety of the flavin [23, 46]. The solvation of the DNA binding pocket suggests that solvent polaron formation and non-equilibrium effects might be involved in the mechanism which underlies the electron transfer from the excited $FADH^{-}$ to the CPD.

During the photo-repair, the excess electron is transferred from the LUMO orbital of the flavin

to the LUMO orbital of the thymine dimer and needs to be treated explicitly. It is straightforward to extend the direct simulation approach and to describe the electron wavefunction as linear combination of LUMO orbitals of the molecular entities involved in the electron transfer. However, careful benchmarking against higher-level methods is required to validate the DFTB-FO approximation used in the direct electron transfer methodology.

So far, the flavin excited state has only been accounted for implicitly in the electron transfer simulations via point charges. Here, inclusion of the flavin excited state in the quantum description is crucial for the electron transfer dynamics and an approximate, yet sufficiently accurate description of the flavin excited state is developed and validated in the following.

6.1 Simulation Setup

6.1.1 MD Simulation

The protein structure of the photolyase-DNA complex used in this study is based on the X-ray crystal structure by Mees *et al.* (PDB ID 1TEZ) [17]. The FAD cofactor was built from riboflavin and adenosine diphosphate (ADP) models using a force field parametrization developed earlier [130, 132]. Force field parameters for the CPD analog were developed based on the CPD structure extracted from PDB entry 1SNH [171]. Gaff [130] force field parameters were used for 8-HDF. Atomic charges for the $FADH^{\bullet}$ and the $FADH^{-}$ as well as the CPD-analog and the HDF-cofactor were determined from RESP calculations [133] and are provided in the Appendix.

The protein model was solvated in a cubic periodic box with 113 Å edge length filled with TIP3P water molecules [149]; neutralized with Na^{+} ions. All simulations were performed with the GROMACS package [150]. The improved Amber99 force field was used [130, 131] for the protein and DNA parts. The system was first energy minimized. Then gradually heated to 300 K during an interval of 100 ps with harmonic constraints imposed on all non-hydrogen atoms except water molecules. After heating the system was equilibrated for 1.4 ns using the Nose-Hoover thermostat, allowing the system to acquire a constant volume at 1 atm pressure. With restraints turned off, a 50 ns MD simulation was carried out. The integration time step was 2 fs. The temperature was kept at 300 K by applying the Nose-Hoover thermostat [101] while the pressure was controlled by the Parrinello-Rahman barostat [102] at a value of 1 atm. The LINCS algorithm [134] was employed to keep bonds at a fixed length. The time evolution of the photolyase-DNA complex backbone root mean square deviation (RMSD) is shown in the appendix A, confirming the stability of the simulation.

6.1.2 Quantum Chemistry Calculations

Quantum chemistry calculations to obtain electron affinities of FAD and CPD were carried out with Gaussian09 [172]. Only the riboflavin part of the FAD terminated by a methyl-group was considered in the quantum chemical description. For the CPD, the sugar parts were replaced by methyl groups. Geometries were optimized at the B3LYP/TZVP level. Single point calculations were performed with HF and DFT methods. The single point energies were computed with the 6-311G** basis set [173] for HF, B3LYP and PBE. Single point energies computed with the long-range and dispersion corrected ωB97XD [89] functional used the cc-pVTZ basis set [174]. Single point calculations with the density functional tight binding (DFTB) method were carried out with the dftb+ [164] program package using the 3OB parameterization [93].

Excited state energies were computed at the TDDFT/B3LYP level employing the 6-31G* basis set with Gaussian09.

The effect of protein and DNA electrostatic potential on the electron affinities was considered by including a distribution of point charges in the quantum chemistry calculations. The set of point charges was obtained for 1000 snapshots of the MD simulation of the photolyase inactive state. Single point calculations in the presence of external point charges were performed at the ωB97XD/cc-pVTZ level of theory using Gaussian09. The same applies for excited state energies of FAD computed with TDDFT/B3LYP employing the 6-31G* basis set. DFTB single point energies in the presence of point charges were computed with the QM/MM interface implemented an in-house version of GROMACS 4.6, using DFTB2 [154] for the QM part.

6.2 Electron vs Hole Transfer

The direct electron transfer approach as described in Sec. 2.3.3, has been developed for the simulation of hole transfer. The energetics of hole transfer are characterized by the energy which is required to remove an electron from the electron transfer site and to create a positive hole. This energy is the ionization potential of the molecule which is directly connected to the HOMO energy via Koopmans' theorem. The hole transfer mechanism using the example of *E. coli* photolyase photo-activation is illustrated in Fig. 3.2.

During the photo-repair of DNA by photolyase, an excess electron is transferred from the LUMO of the flavin to the LUMO of the thymine dimer as shown in Fig. 6.1. Here, the energetics of electron transfer are determined by the energy which is gained by attaching an additional electron to the electron transfer site. This energy is the electron affinity of the molecule which is directly connected to the LUMO energy via Koopmans' theorem.

The calculation of electron affinities has proven more challenging than the calculation of ionization potentials [76]. Due to the more diffuse electron charge distribution of the anions com-

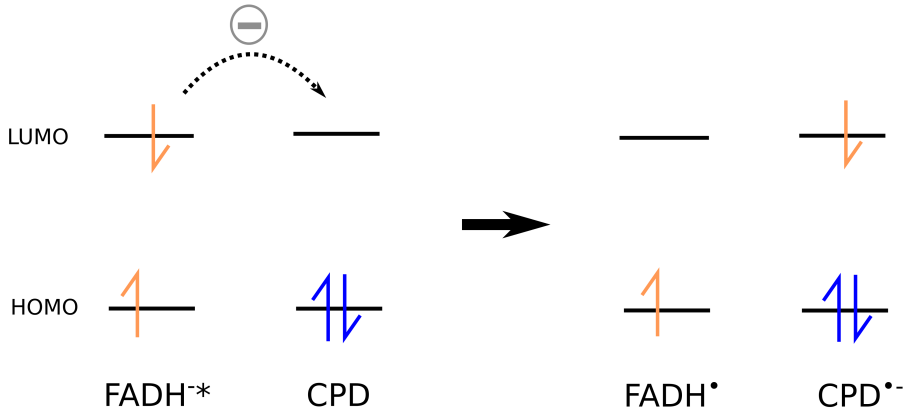


Figure 6.1: Initial electron transfer during photo-repair. *The excess electron is transferred from the LUMO orbital of the flavin to the LUMO orbital of the thymine dimer. The direct electron transfer approach can easily be extended but the representation of the electron transfer energetics via LUMO energies in the DFTB-FO parameterization needs to be validated. An approximate, yet sufficiently accurate description of the flavin excited state needs to be developed and tested.*

pared to cations, electron affinities are found to be more sensitive to the choice of basis set used than ionization potentials [76, 175]. The diffuse electron charge density requires the use of computationally expensive, large basis sets with diffuse functions.

Furthermore, the excess electron might be not or too weakly bound due to the DFT self-interaction error [176]. This effect is found to be less pronounced for larger molecules where the electron is more delocalized. Furthermore, due to the use of finite basis sets, the additional electron is forced to remain within the spatial area defined by basis functions, which counterbalances the self-interaction based error [76]. In practice, the electron affinities calculated with DFT methods for a large set of molecules yield excellent results, for example using the B3LYP functional, electron affinities with overall mean errors of 0.2 eV were reported [177].

As described in the previous chapter 5, Koopmans' theorem (Eq. 5.1) relates the ionization potential to the HOMO energy of a molecule. The relation is also true for the electron affinity and the energy of the lowest unoccupied molecular orbital (LUMO) for HF theory:

$$EA = E^N - E^{N+1} = -\Delta E_R = -\epsilon_L \tag{6.1}$$

A similar theorem exists for DFT and is referred to as Janak's Theorem. It is important to note that Janaks' theorem only holds for the correct exchange correlation functional. As the exchange-correlation functional is always approximated, the orbital energies calculated with DFT are expected to deviate from the actual ionization potential/electron affinity and depend on the func-

tional used.

The negative of the DFT HOMO energy is typically found to be too small compared to the ionization potential, still it was shown to be related to the experimental ionization potential by a constant shift [178]. A similar relationship was also found for HOMO energies and ionization potential of purine nucleobases calculated with the DFTB method and was successfully applied to simulate hole transfer in DNA [82].

The relation between LUMO energies and electron affinities seems to be less clear. For molecules with bound anionic states, a linear correlation was found between the DFT LUMO energy and the calculated and experimental electron affinities. On the contrary, for molecules with unbound anionic states, the same study did not find a satisfactory linear correlation even when a larger basis set was employed [175].

In terms of methodology, it is straightforward to extend the direct simulation approach and to describe the excess electron wavefunction as a linear combination of LUMO orbitals of the molecular entities involved in the electron transfer. However, the basis set dependence of the electron affinities of flavin and thymine dimer needs to be tested carefully beforehand and the suitability of the LUMO energies as estimate of the electron affinity for these molecules must be investigated. In addition, careful benchmarking of the DFTB LUMO energies against electron affinities yielded by higher-level methods is required to validate the DFTB-FO approximation used in the direct electron transfer approach.

6.3 Relative Energies of Flavin and Thymine Dimer States

The first step is to benchmark DFTB against DFT and ab initio methods. The vertical relative energies can be obtained using the ΔSCF procedure which involves the calculation of the total energy of both the neutral and the charged molecule based on the geometry of the neutral state.

Table 6.1: Vertical relative energies for flavin and thymin dimer in different states. *Calculated with DFT and HF methods. Energies are given in eV.*

	DFTB	PBE	B3LYP	HF	ω97XD	Exp.
$FADH^-$	0	0	0	0	0	
$FADH^\bullet$	2.41	2.08	2.23	1.43	2.43	
$FADH^{-\star}$	-	-	2.86	-	3.4	2.49
CPD	0	0	0	0	0	
$CPD^{\bullet-}$	1.06	0.47	0.64	2.52	0.88	

The relative energies of different states of flavin and CPD are presented in Tab. 6.1. $FADH^\bullet$ is found to be higher in energy than $FADH^-$ for all methods and functionals used. This means

that the vertical ionization potential of $FADH^-$ is positive. The first singlet excited state of flavin $FADH^{-*}$ was calculated with TDDFT using two different functionals and was found to be higher in energy than $FADH^\bullet$. While the excitation energy obtained with the B3LYP functional matches the experimental value [179, 180] and higher level methods surprisingly well (see appendix C), the ω97XD functional overestimates it by about 1 eV.

The relative energy of $FADH^\bullet$ obtained with DFTB agrees quite well with the relative energy obtained with the ω97XD functional. The B3LYP and PBE functionals yield slightly lower relative energies, while the relative energy calculated with HF is significantly lower compared to the DFT methods.

The relative energy of the anionic thymine dimer radical is found to be higher in energy than in the neutral state, indicating a negative electron affinity for CPD. However, the exact value of the relative energy seems to depend on the method used. DFTB yields a relative energy of about 1.1 eV and is in reasonable agreement with the relative energy computed using the ω97XD functional. The B3LYP and PBE functionals yield lower relative energies, deviating by up to 0.4 eV from the value obtained with the ω97XD functional. The relative energy obtained with HF is significantly larger compared to all other methods.

6.4 Influence of Diffuse Functions on the Relative Energies

In an attempt to determine the relative energy of the thymine dimer more accurately, the calculations above were repeated with larger basis sets including diffuse functions. For PBE, B3LYP and HF the 6-311++G** basis set was employed, for the ω97XD functional the aug-ccpVTZ basis set was used. The results are presented in Tab. 6.2.

Table 6.2: Effect of diffuse functions on the vertical relative energies. *Calculated with DFT and HF methods using basis sets augmented with diffuse functions. Energies are given in eV.*

	DFTB	PBE	B3LYP	HF	ω97XD
$FADH^-$	0	0	0	0	0
$FADH^\bullet$	-	2.36	2.50	1.67	-
CPD	0	0	0	0	0
$CPD^{\bullet-}$	-	-0.23	-0.09	0.36	0.06

The effect of diffuse functions on the relative energy of CPD is apparent. The relative energies are significantly lower compared to the calculations where basis sets without diffuse functions were used. At the B3LYP and PBE level of theory the relative energy is found to be negative, which in turn means the electron affinity of CPD is positive, indicating a stable anionic CPD radical.

However, the absolute value of the electron affinity of CPD still depends on the method and functional used. The slightly negative electron affinity obtained with the ω97XD functional is in line with the results of a recent benchmarking study [181]. For flavin, the effect of diffuse functions is less pronounced and the relative energies are only slightly higher compared to computations where no diffuse functions were used.

It was shown that the determination of gas-phase electron affinities of nucleic acid bases is a difficult problem which is partly due to the formation of dipole bound anions of these system [182]. A dipole bound state is stabilized primarily due to electrostatic interaction between the electron and the dipole moment of the molecule. If the dipole moment of a molecule is larger than 2.5 D, an excess electron can be trapped in a dipole bound state [183] so that the excess electron is not bound in the valence orbitals close to the nuclear framework but rather in an extended, diffuse orbital off the molecular framework. However, it was demonstrated that the dipole bound states in the thymine anion can be obtained with the use of diffuse functions and taking into account electron correlation effects [182]. Both flavin and the thymine dimer are found to have sufficiently high dipole moments to form dipole bound states of more than 10 D and 6 D, respectively. The dipole moments obtained for different states of flavin and the thymine dimer are listed in Tab. D in the appendix. A positive electron affinity for CPD is only obtained when diffuse functions are included in the basis set. This indicates, that with the use of diffuse functions the electron affinity in the dipole bound state of the anionic CPD radical is obtained.

Thymine dipole bound anions are observed experimentally only for free thymine. Upon addition of a water molecule the valence bound anion was detected [184]. A recent benchmarking study computed electron affinities for CPD at the ω97XD/aug-cc-TZVP level of theory [181] and found the electron affinity to be -0.06 eV. Upon inclusion of the sugar moieties a higher electron affinity was obtained and the effect of the diffuse functions seems to be less pronounced. With diffuse functions the electron affinity was found to be -0.69 eV, without diffuse functions a value of -0.98 eV was obtained. It is therefore concluded that dipole bound state are negligible if protein and solvent environment are included in the calculation and that smaller, computationally less expensive basis sets without diffuse functions are sufficient for the calculation of the electron affinities here.

6.5 Model of the Excited State in the DFTB-FO Parameterization

The relevant states of flavin and CPD are sketched in the energy level diagram shown in Fig: 6.2. $FADH^{\bullet}$ is higher in energy than $FADH^{-}$ but lower in energy than the singlet excited state $FADH^{-*}$. Upon electron transfer to the CPD, the flavin is converted from its singlet excited state to the neutral semiquinone state. The energy which is released during this reaction is indicated by the blue

arrow. This energy might be referred to as ionization potential of the flavin singlet excited state $FADH^{-*}$. The anionic $CPD^{\bullet-}$ is higher in energy than in its neutral state. The energy which is required to convert neutral CPD to anionic $CPD^{\bullet-}$ is described by its electron affinity. The difference of the ionization potential of the flavin singlet excited state $FADH^{-*}$ and the CPD electron affinity defines the driving force of the electron transfer reaction.

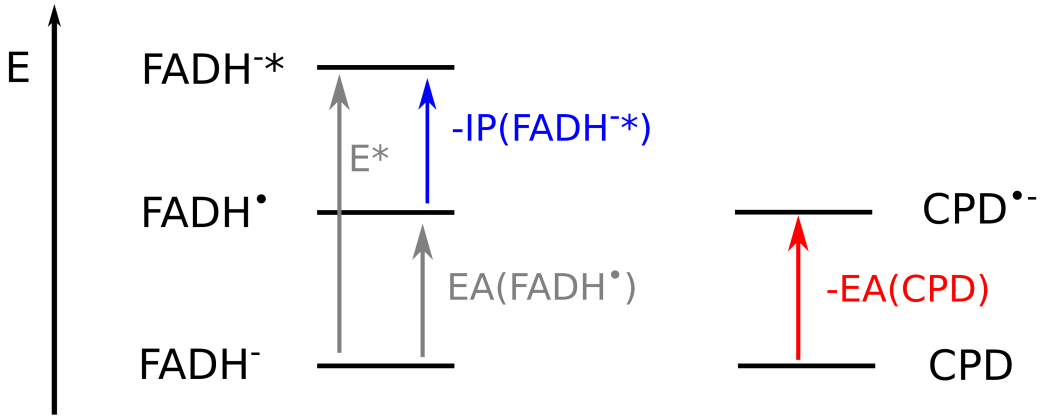


Figure 6.2: Energy level diagram for flavin and CPD. *The blue arrow indicates the ionization potential of $FADH^{*-}$ and the red arrow denotes the electron affinity of CPD. Their difference represents the driving force for the electron transfer reaction. The ionization potential of $FADH^{*-}$ can be calculated from the singlet excitation energy E* and the $FADH^{\bullet}$ electron affinity (grey).*

The relative energies obtained from gas-phase calculations can be used to estimate the driving force for the reaction

$$FADH^{*-} + CPD \rightarrow FADH^{\bullet} + CPD^{\bullet-} \tag{6.2}$$

and can be obtained as:

$$\Delta E = E(FADH^{\bullet}) + E(CPD^{\bullet-}) - E(FADH^{*-}) - E(CPD). \tag{6.3}$$

Ultimately, the aim of the DFTB-FO parameterization is to describe the energetics of the electron transfer reaction via the LUMO energies of the molecules. This is straightforward for CPD. The energy required to convert neutral CPD to anionic $CPD^{\bullet-}$ is described by its electron affinity and therefore directly related to the CPD LUMO energy via Koopmans' theorem. The ionization energy of the flavin singlet excited state, on the other hand, cannot be connected to the LUMO

energy of $FADH^-$ directly. Therefore an alternative description is needed. The ionization energy of the flavin singlet excited state can be obtained as the difference of the vertical singlet excitation energy and the electron affinity of the flavin in its neutral semiquinone state(see Fig. 6.2). Rearrangement of the terms in Eq. 6.3 yields an expression for the reaction energy which includes the ΔSCF energies of flavin and CPD and the energy of the first excited state of $FADH^{*-}$:

$$\Delta E = E(CPD \rightarrow CPD^{\bullet -}) + E(FADH^- \rightarrow FADH^{\bullet}) - E* \tag{6.4}$$

$$= -EA(CPD) + EA(FADH^{\bullet}) - E* \tag{6.5}$$

In the DFTB-FO parameterization the ionization energy of $FADH^{*-}$ can be computed as the $FADH^{\bullet}$ LUMO energy shifted by the vertical singlet excitation energy. If this approximation is a valid description of the ionization energy of the flavin singlet excited state and if the $FADH^{\bullet}$ LUMO energy in the DFTB-FO parameterization is a good estimate of its electron affinity will be tested in the following.

6.6 Test of the Electron Affinities in the DFTB-FO Parameterization

The energetics of electron transfer can be determined from the ionization potential of $FADH^{*-}$ and the electron affinity of CPD. As outlined above, the ionization potential of $FADH^{*-}$ can be approximated by difference of the electron affinity of the flavin in its neutral semiquinone state $FADH^{\bullet}$ and the singlet excitation energy of $FADH^-$. The electron affinity can be computed by means of the ΔSCF procedure as the difference of the total energy of the neutral and the anionic state of a molecule. A more efficient way to estimate the electron affinity of a molecule is the calculation of the LUMO energy ϵ_L.

Table 6.3 compares the electron affinities obtained as LUMO energy to those calculated with several DFT methods, HF and the ω97XD functional using the ΔSCF procedure. Note that the values for ϵ_L and ΔSCF-v, shown in Table 6.3, equal the negative of the electron affinity according to Eq. 6.1. The LUMO energy of $FADH^{\bullet}$ calculated with DFTB and PBE matches quite well the vertical electron affinity obtained by means of the ΔSCF procedure with the respective method. The LUMO energy obtained with B3LYP differs from the ΔSCF value by about 0.5 eV. For HF and the ω97XD functional, the deviation exceeds more than 2 eV. The positive LUMO energies obtained with the two latter methods suggest that the excess electron is not bound. This is in contrast to the positive electron affinity of $FADH^{\bullet}$ which was found for all methods.

In the case of CPD, the LUMO energy agrees well with the ΔSCF values only for the ω97XD functional. The energies computed with HF deviate by about 0.6 eV, while the LUMO energies cal-

culated with the other methods differ by more than 1.7 eV from the electron affinity obtained by means of the ΔSCF procedure. The electron affinity of CPD is negative with all methods. This is in agreement with the positive LUMO energy of CPD calculated with HF and the ω97XD functional which indicates that the excess electron is not bound.

The data presented here do not provide a clear picture of the electron affinities of flavin and CPD. However, the long-range dispersion corrected ω97XD functional is the most sophisticated of the functionals and methods considered here and is expected to provide the best results for the electron affinities obtained with the ΔSCF procedure. For $FADH^{\bullet}$, the DFTB LUMO energy agrees reasonably well with the ΔSCF energy obtained with the ω97XD functional. The CPD LUMO energy is severely underestimated by DFTB compared to the ΔSCF energy obtained with the ω97XD functional.

Table 6.3: Electron affinities of $FADH^{\bullet}$ and CPD as ΔSCF and $\epsilon_{\mathbf{LUMO}}$. *DFT and HF methods (6-311G** basis set) and the ω97XD functional (cc-pVTZ) are used.*

	$FADH^{\bullet}$		CPD		relative EA	
method	ϵ_{LUMO}	ΔSCF-v	ϵ_{LUMO}	ΔSCF-v	ΔSCF-v	ϵ_{LUMO}
DFTB	-1.96	-1.90	-1.44	1.06	-1.96	-0.52
PBE	-1.83	-1.62	-1.81	0.47	-2.09	-0.01
B3LYP	-1.18	-1.69	-1.06	0.63	-2.23	-0.12
HF	3.04	-0.62	3.10	2.48	-3.10	0.56
ω97XD	0.62	-1.77	0.91	0.88	-2.65	-0.26

The difference between the electron affinity of $FADH^{\bullet}$ and the electron affinity of CPD, the relative electron affinity, is the quantity that determines the energetics of electron transfer. The CPD LUMO energy calculated with DFTB can be easily corrected by adding a constant value so that the correct energy difference between the $FADH^{\bullet}$ LUMO and the CPD LUMO is obtained. The difference between the ΔSCF energies obtained with the ω97XD functional might be used as a reference. However, it is important to verify that the DFTB LUMO energies show identical behavior to the values obtained with ω97XD functional by the ΔSCF procedure when computed along molecular dynamics trajectories. This will be thoroughly tested in the following section.

6.7 Effect of Environment and Structural Variation on the DFTB-FO Parameters

In order to describe the electron transfer energetics correctly, it is crucial that the electron affinities obtained as DFTB LUMO energies reproduce the time evolution and structural fluctuation of the electron affinities calculated by the means of the ΔSCF with a higher level method. Therefore, the electron affinities are evaluated along molecular dynamics trajectories. Furthermore, the

electrostatic effect of the environment on the flavin singlet excitation energy is considered.

The effect of protein and solvent electrostatic potential on the relative energies is accounted for by including a distribution of point charges in the quantum chemistry calculations. The geometries of flavin and CPD and the set of point charges are taken from a MD simulation of the resting state of photolyase in complex with DNA. The electron affinities and the flavin singlet excitation energy are calculated in the presence of point charges for 100 snapshots taken at 1 fs intervals.

The electron affinities are calculated as LUMO energies using DFTB and and by the means of the ΔSCF procedure employing the ω97XD functional, which was chosen as a reference based on the discussion in the previous sections.

For the energetics of electron transfer, computation of the ionization potential of $FADH^{-*}$ is required. According to the model presented in chapter 6.5, the $FADH^{-*}$ ionization potential is calculated as the difference of the vertical singlet excitation energy of $FADH^{-}$ and the $FADH^{\bullet}$ electron affinity. Fig. 6.4A shows the time evolution of the electron affinities of $FADH^{\bullet}$ and CPD calculated with the ω97XD functional using the ΔSCF procedure, the excitation energy of $FADH^{-}$ obtained with TDDFT as well as the $FADH^{-*}$ ionization potential.

In principle, the true $FADH^{-*}$ ionization potential can be represented either by the LUMO of the $FADH^{-}$, the LUMO of the $FADH^{\bullet}$ or the SOMO of the $FADH^{-}$. The SOMO is the highest, single occupied molecular orbital and, in the case of flavin, is the same orbital as the LUMO of the other flavin states but occupied by a single electron. The dominant electronic configurations of the three different flavin states are shown in Fig. 6.3.

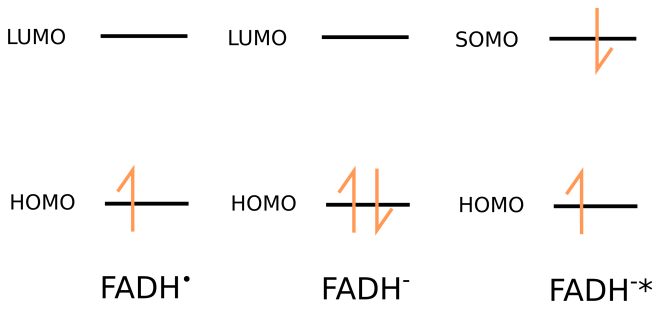


Figure 6.3: Electronic configurations of different flavin states. *The true $FADH^{-*}$ ionization potential can be represented either by the LUMO of $FADH^{-}$, the LUMO of $FADH^{\bullet}$ or the SOMO of $FADH^{-}$. The SOMO is the highest, single occupied molecular orbital.*

The total energy of the charged system in the electron transfer approach is obtained by expanding the electron density around the reference density of the neutral system. In this spirit, the LUMO of $FADH^{\bullet}$ would be the quantity of choice. Still, in the following the DFTB LUMO energies of all three possible electronic configurations are compared to the true $FADH^{-*}$ ionization

potential obtained with ω97XD functional and TDDFT in order to find the best representation of $FADH^{-*}$. The time evolution of the energies obtained with DFTB are shown in Fig. 6.4B.

Time evolution and fluctuations of the $FADH^{\bullet}$ LUMO energy computed with DFTB agree excellently with the $FADH^{-*}$ ionization potential obtained with ω97XD functional and TDDFT. The CPD LUMO energy calculated with DFTB shows a very similar time evolution and follows the fluctuations of the electron affinity obtained with the ω97XD functional in a reasonable agreement. The fluctuations of the CPD LUMO energy seem to be somewhat more pronounced than the fluctuations of the CPD ΔSCF energy.

The $FADH^{-*}$ ionization potential and the CPD electron affinity obtained with the ω97XD functional and TDDFT is underestimated by the DFTB LUMO energy for both molecules by almost 2.5 eV.

The energies of the $FADH^{-}$ LUMO and SOMO deviate by about 2 eV from the $FADH^{-*}$ ionization potential but follow the dynamical fluctuations quite accurately. The fluctuations of the LUMO of the $FADH^{\bullet}$ state and the SOMO of the $FADH^{-}$ state seem to agree slightly better with the $FADH^{-*}$ ionization potential than the fluctuations of the LUMO of the $FADH^{-}$ state. The dynamic features of the DFTB LUMOs agree nicely with those of the ω97XD functional relative energies. Since the relative fluctuations are the crucial quantity for the description of electron transfer, the DFTB LUMO energies are well suited to describe the energetics.

In order to determine how the different LUMOs reproduce the structural fluctuations of the $FADH^{-*}$ ionization potential and CPD electron affinity and to quantify the exact deviation of the LUMO energies, the quantum chemical calculations were carried out for a larger number of structures. The complete ensemble of structural variety of the photolyase/DNA complex is taken into account by carrying out quantum chemical calculations for 1000 structures and charges taken from snapshots of the MD simulation. As outlined in chapter 5.3, the deviation of the DFTB LUMO energy from the $FADH^{-*}$ ionization potential and CPD electron affinity, respectively, by a constant C can be determined by minimizing the deviation function Eq. 5.2. The results of the numerical fit for CPD and the different flavin states are summarized in Tab. 6.4.

Table 6.4: Deviation of the DFTB LUMO energies from the true IP/EA values obtained with ω97XD. *Deviation constant and RMSD are calculated for different flavin states and CPD. Energies are given in eV.*

	deviation constant C	RMSD
$FADH^{\bullet}$LUMO	-2.526	0.028
$FADH^{-}$LUMO	+1.911	0.032
$FADH^{-}$SOMO	+2.047	0.028
CPD LUMO	-2.291	0.139

The structural fluctuations of the $FADH^{-*}$ ionization potential are well reproduced by the DFTB

LUMOs of all three flavin electronic states, with the $FADH^{\bullet}$ LUMO and the $FADH^{-}$ SOMO showing a slightly smaller RMSD of 0.028 eV compared to the $FADH^{-}$ LUMO with an RMSD of 0.032 eV. The DFTB LUMO of CPD reproduces the structural fluctuations of the ΔSCF electron affinity up to an reasonable accuracy of 0.139 eV. As shown by a recent quantum chemical study, the electron affinity seems to be sensitive to the CPD geometry [181] and as pointed out above, depends on the functional and basis set used. Correlation effects were suggested to be important for the correct description of the CPD dipole bound states. These effects might be better accounted for at the ω97XD/cc-pVTZ level of theory than by DFTB2. Altogether, the dynamic features of the DFTB LUMOs agree well with those of the ω97XD functional relative energies and are suited to describe the energetics. Therefore, an additive shift can be safely applied to ensure correct relative energies.

The deviation constants in Tab. 6.4 indicate that the LUMO energies need to be modified by application of an additive shift in order to obtain the correct relative energy of the flavin and CPD DFTB LUMO energies. For example, a shift of 0.235 eV needs to be added to the $FADH^{\bullet}$ LUMO energy in order to obtain the correct energy difference compared to the CPD LUMO.

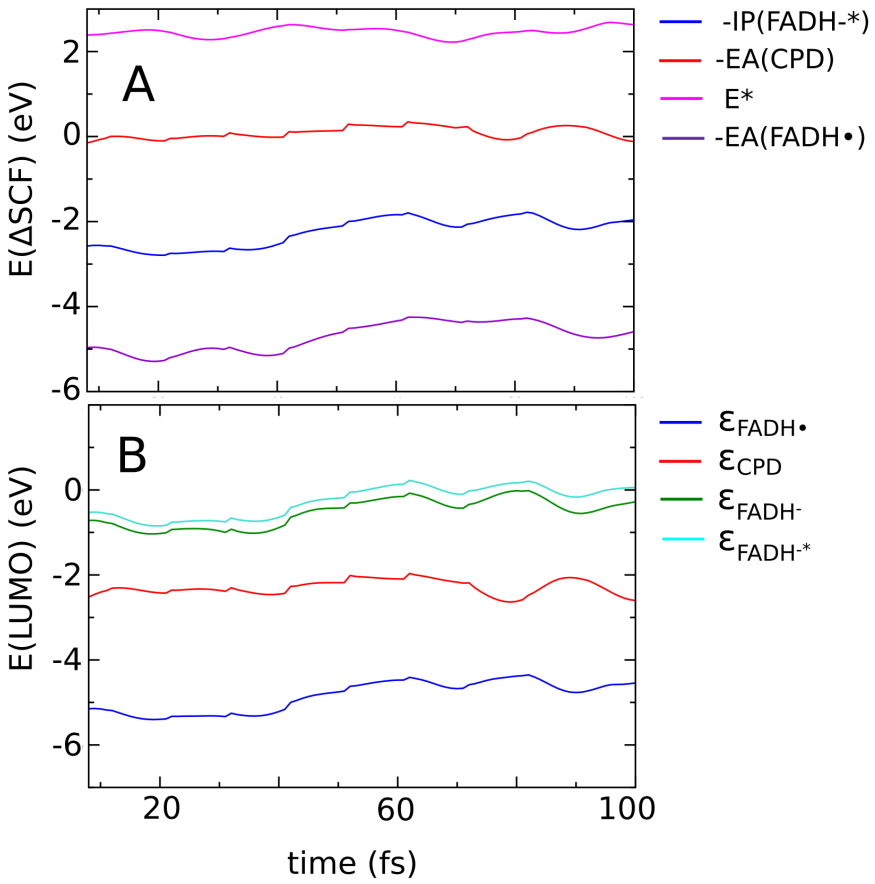


Figure 6.4: Comparison of the energies calculated as ΔSCF and DFTB LUMO energy *for a 100 fs molecular dynamics trajectory.* ***A*** *The $FADH^{-*}$ ionization potential (blue) is calculated as the difference of the flavin singlet excitation energy E* (magenta) and the $FADH^{\bullet}$ electron affinity.* ***B*** *The DFTB LUMO energy of the different flavin states $FADH^{\bullet}$ (blue), $FADH^{-}$ (green) and $FADH^{-*}$ (mint) follow the time evolution of the $FADH^{-*}$ ionization potential obtained by means of the ΔSCF procedure quite accurately, the CPD LUMO energy obtained with DFTB follows the structural fluctuations of CPD electron affinity in reasonable agreement. The DFTB LUMO energies are therefore suited to describe the electron transfer energetics.*

6.8 Inner-sphere Reorganization Energy

The inner-sphere reorganization energy is considered as effective parameter in the direct electron transfer approach. Due to the force field treatment, the equilibrium geometry of a molecular fragment does not change upon charging. A feasable approach to describe this phenomenon is to introduce a new contribution to the total energy which is a quadratic function of the excess charge ΔQ_m:

$$E_{\lambda_i} = \epsilon_m - \lambda_i^{(m)} \Delta Q_m^2. \tag{6.6}$$

The inner sphere contribution to the reorganization energy, λ_i, of fragment F must be determined for each new fragment. λ_i is obtained from *ab initio* calculations

$$\lambda_i(F) = E_{F^-}^{QM}(F) - E_F^{QM}(F) \tag{6.7}$$

where E^{QM} denotes the single point energy of the neutral fragment (F). Single point energy calculations are carried out twice, using the optimized geometry of the neutral fragment F and the optimized geometry of the charged fragment F^-. The subscripts denote the charge state of the fragment during the geometry optimization. CPD geometry optimizations were carried out at the ω97XD/cc-pVTZ level of theory using Gaussian09 [172]. The optimized geometries of the $FADH^{-*}$ and $FADH^{\bullet}$ states were obtained with the TurboMole 5.1 program package [185] using the B3LYP functional and the def2-TZVP basis set[186]. The single point energies were evaluated with Gaussian09 at the ω97XD/cc-pVTZ level of theory using the COSMO implicit solvent model [187] for both geometries of each molecule. Inner reorganization energies of 1.38 eV and 0.31 eV were obtained for CPD and $FADH^{\bullet}$, respectively. The value obtained for CPD is strikingly large. This might be attributed to the sensitivity of the CPD energy on its geometry as pointed out in [181].

6.9 Differential Charges

The fragment charges obtained after each TDDFT calculation step in the direct electron transfer scheme are projected to MM atomic charges. These atomic charges are used in the subsequent molecular dynamics step and thereby induce a polarization of the environment by the moving charge. This effect is crucial for the understanding of the electron transfer process.

During the simulation, the atomic charges are determined by interpolating between the atomic charges of the neutral and the charged fragment according to the fraction of charge on the fragment at a given time instant. The atomic charges of both the neutral and the charged molecule are tabulated and need to be obtained beforehand by RESP calculations. The atomic charges of different states of flavin and CPD were determined in this study and are given in the Appendix E. The molecular dipole moment obtained from the quantum chemical calculation is well reproduced

by the RESP atomic charges as shown in Tab 6.5.

Table 6.5: Dipole moments of flavin and CPD in the ground state. *The dipole moments obtained from quantum chemical calculations are well reproduced by the RESP charges. Dipole moments are given in debeye.*

redox state	B3LYP/631G*	RESP
CPD	7.5	7.94
CPD-	9.8	9.62
$FADH^{\bullet}$	8.16	8.2
$FADH^{-}$	11.23	11.11

The molecular dipole moment of CPD and $FADH^{\bullet}$ increases upon charging by about 2.3 and 3 D, respectively. Determination of the atomic charges of the flavin singlet excited state $FADH^{-*}$ is more involved. The molecular dipole moment of $FADH^{-}$ was reported to change upon excitation, however, results from different theoretical studies and experiment were ambiguous. A recent theoretical study reported a difference of about 4 D between the ground and excited state dipole moment of the flavin molecule [35]. Based on the results of time-resolved femtosecond spectroscopy measurements the flavin molecular dipole moment was estimated to vary by about 1-2 eV between ground and excited state [188].

The molecular dipole moment of the first singlet excited state was calculated at different levels of theory in this study. The results are listed in Tab. 6.6.

Table 6.6: Dipole moment of the first singlet excited state of flavin. *Obtained at different levels of theory. Dipole moments are given in debeye.*

electronic state	TDDFT(B3LYP-631G*)	CASPT2 [189]	SORCI (HF/def2-TZVP)
S0	10.75	12.25	11.68
S1	5.14	6.61	5.39

The difference between the molecular dipole moments of the ground and first excited state of $FADH^{-}$ obtained with TDDFT is in agreement with results from CASPT2 and SORCI. The molecular dipole moment of $FADH^{-*}$ is 5.6 D smaller than that of $FADH^{-}$. The absolute molecular dipole moments are slightly underestimated by TDDFT compared to the higher level methods.

The findings suggest that the molecular dipole moment of the flavin molecule will change several times during the photo-repair reaction. Upon light excitation, it decreases by 5.6 D and subsequently increase by 3 D once the electron is transferred. The changing dipole moment will result in dynamic repolarization of the environment and might affect the electron transfer dynamics substantially.

6.10 Discussion

For the simulation of photo-repair by CPD photolyase, the description of the electron transfer energetics via DFTB LUMO orbitals needs to be tested, and an approximate, yet sufficiently accurate description of the flavin excited state needs to be developed. The first step is to benchmark DFTB against higher level methods. The energies obtained from the gas phase calculations show that flavin is well described by DFTB compared to the long-range dispersion corrected ω97XD functional. Also, the energy of the flavin singlet excited state computed at the TDDFT/B3LYP level of theory is in a good agreement with the available experimental data.

The results for the thymine dimer are less clear. A negative electron affinity is found, indicating that the CPD radical anion is not stable, but the magnitude of the electron affinity varies depending on the method and basis set used. However, the relative DFTB energies are found in a reasonable agreement with the results obtained with the ω97XD functional. In an attempt to determine the electron affinity of CPD more accurately, basis sets augmented with diffuse functions are used. While the energies of the different flavin states change little when diffuse functions are included, their effect of the thymine dimer energies is apparent. Its electron affinity becomes slightly positive, indicating a stable dipole-bound state. However, the absolute values still vary depending on the method and basis set used. It was shown that dipole-bound states of thymine radical anions only exist in gas-phase and disappear if water molecules are included [182]. It is therefore concluded that dipole-bound state are negligible if protein and solvent environment are considered in the calculation and that smaller, computationally less expensive basis sets without diffuse functions are sufficient for the calculation of the electron affinities here.

The goal in the development of the model parameters is to describe the energetics of the electron transfer reaction via the LUMO energies of the molecules. The LUMO energy of CPD is connected directly to its electron affinity via Koopmans' theorem. The ionization potential of an excited state, on the other hand, cannot be represented directly by the orbital energy. Still, the LUMO energy of $FADH^{\bullet}$, which is linked to its electron affinity, might be used, as long as a shift by the vertical singlet excitation energy is applied.

The LUMO energy computed with DFTB is a good estimate of the $FADH^{\bullet}$ electron affinity obtained with the ω97XD functional by means of the ΔSCF procedure. The LUMO energies of the thymine dimer computed with DFT methods overestimate the electron affinity obtained with the ω97XD functional severely. Consequently, the difference between the $FADH^{\bullet}$ and CPD electron affinities, which determines the energetics of electron transfer, is not well approximated by the DFTB LUMO energies. In principle, the LUMO energies can be modified in order to obtain the correct energy difference. However, it must be ensured that the dynamical features of the DFTB LUMOs agree with those of the ω97XD functional relative energies, especially in the case of CPD.

The time evolution and fluctuations of the $FADH^{\bullet}$ LUMO energy computed with DFTB along

a molecular dynamics trajectory agree excellently with those of the $FADH^{-*}$ ionization potential obtained as difference of the ω97XD $FADH^{\bullet}$ electron affinity and the TDDFT singlet excited state energy. The $FADH^{-}$ LUMO and $FADH^{-}$ SOMO energies are also considered, but their fluctuations agree slightly less perfect with the true $FADH^{-*}$ ionization potential than those of the $FADH^{\bullet}$ LUMO. The CPD LUMO energy shows a very similar time evolution and follows the fluctuations of the electron affinity obtained with the ω97XD functional, in reasonably well agreement. Importantly, this indicates that the DFTB LUMO energies are suited to describe the energetics of both of the flavin singlet excited state and of the thymine dimer. Therefore, the correct energy difference between $FADH^{-*}$ and CPD is obtained by applying an additive shift to the LUMO energy of the molecules.

The gas-phase calculations suggest that the electron transfer is slightly endothermic. The energies obtained with a QM/MM treatment of the protein and solvent environment indicate that the electron transfer becomes even more endothermic by about 2.5 eV. Note that, the QM/MM simulation were performed with protein and solvent environment fully relaxed to the $FADH^{-}$ state, thus stabilizing it. Importantly, the flavin dipole moment decreases by more than 5 D upon light excitation. This indicates that the singlet excited state of flavin will be less stabilized by the environment than the $FADH^{-}$ state. Dynamic re-polarization of the environment might likely bring the energy level of flavin closer to that of CPD, so that the population of $CPD^{\bullet-}$ becomes more favorable than population of $FADH^{-*}$. Dynamic re-polarization of the environment was found to provide the driving force for electron transfer during the photo-activation of cryptochrome and photolyase [126, 146, 158] and this effect might be crucial for the DNA repair as well.

Dynamic electron transfer simulations need to be carried out to test this hypothesis and investigate the molecular mechanism of the electron transfer dynamics in detail. Comparison of the results with existing experimental data will allow for validation of the model. The model parametrization based on the DFTB LUMO energies developed above on the basis of results yielded by higher level methods provides the necessary prerequisites to carry out direct electron transfer dynamic simulations, in which the dynamic interaction of the moving electron with the environmental fluctuations is fully accounted for.

CHAPTER 7

Conclusion

In this work, light-induced charge transfer reactions in proteins of the photolyase/cryptochrome family are studied using approximate QM/MM molecular dynamics schemes. In particular, the approach is applied to the photo-activation of *E. coli* photolyase, which is well investigated and represents the model system of the protein family, and to *A. thal.* cryptochrome whose activation is less well understood. Furthermore, an extension of the simulation approach to electron transfer via LUMO orbitals is tested and a set of system specific model parameters for the simulation of the DNA repair by CPD photolyase is developed and validated.

Irradiation with UV-light damages deoxyribonucleic acid (DNA) and most commonly leads to formation of a cyclobutane pyrimidine dimer (CPD) by covalent binding of neighboring thymine nucleobases [3]. This defect can be repaired by a photolyase enzyme which transfers an excess electron to the CPD and thereby induces breaking of the bond [23, 24]. The closely related cryptochromes mediate light-induced responses by cycling between different redox states of their flavin cofactor. This is achieved by the process of photo-activation, involving a long-range electron transfer along a chain of conserved tryptophan residues [44, 45, 66, 190].

The theoretical investigation of electron charge transfer processes in proteins is challenging as these reactions occur over long time and length scales and the heterogeneous environment of the protein and the solvent. Most studies of electron transfer in biomolecules are based on Marcus' theory [79, 80, 105, 106] which is, however, limited to certain assumptions about the nature of the transfer. Over the last years, an approximate electron transfer approach has been developed in our group which has been shown to reproduce the predictions of high-level methodology [82, 124, 125]. Simultaneous propagation of the electronic wavefunction and the nuclei yields unbiased dynamics and accounts for both the effect of the environmental fluctuations on the electron and the re-polarization of the environment by the moving electron. Simulation of electron transfer

reactions over long timescales and large systems becomes feasible by considering only frontier orbitals and using the semi-empirical DFTB2 method [154] for the calculation of electron transfer parameters.

Initially developed for the electron transfer in DNA, the direct electron transfer approach has also proven to reproduce the important characteristics of the photo-activation reaction in *E. coli* photolyase, suggesting non-equilibrium effects to be involved [126]. However, open questions regarding the molecular mechanism underlying the electron transfer during the photo-activation process remained.

In this work, additional amino acids in close vicinity of the tryptophan triad were considered in the quantum chemical description. It was shown that the main pathway was identified correctly among the alternative routes. Furthermore, a small but significant contribution of a fourth tryptophan to the electron flow was predicted which was supported by time resolved femtosecond spectroscopy measurements [46].

The electronic polarization effects, which were not accounted for due to the force field treatment, were included in an approximate fashion by scaling the electrostatic interaction of the electron with the molecular mechanics environment. Among several scaling factors tested, a value of 1.4, which matches the suggestions by Leontyev and Stuchebrukhov [142], yielded improved electron transfer rates. The absolute values were found the be the same order of magnitude as the experimental values. Most importantly, the ratio between the individual electron transfer steps was reproduced in good qualitative agreement with the experiment. The findings proved the direct electron transfer approach to be a valuable predictive tool to establish electron transfer pathways in proteins where experimental validation is lacking.

Analysis of the residence time of the tryptophan radicals during the electron transfer dynamics showed the third electron transfer step to occur out of non-equilibrium conformations before full environmental relaxation has set in. This explained the high electron transfer rate that could not be reproduced by classical Marcus' theory [126]. The molecular mechanism underlying the non-equilibrium dynamics seemed to originate in the different degrees of solvation of the individual tryptophans. It was concluded that the structure and dynamics of the *E. coli* photolyase have likely evolved to prevent the middle tryptophan from becoming a charge trap state and ensure a fast and efficient photo-activation process.

The photo-activation of *A. thaliana* cryptochrome involves both electron and proton transfers [64]. Direct electron transfer and potential of mean force calculations could add to solving the discrepancy [58, 64, 67] regarding the rates of proton and electron transfer. The results showed the reactions to be separated in time by several orders of magnitude when the complete protein and solvent environment were considered.

Dynamic re-polarization of the solvent by the moving electron was determined as driving force for the electron transfer similar to *E. coli* photolyase. Given the fact that the tryptophan triad is

conserved throughout the cryptochrome/photolyase family, this suggested that the solvent driving force is a general principle governing the fast electron transfer process of the photo-activation reaction. It was also concluded that it represents the main feature that allows the formation of a radical pair, comprising all the essential prerequisites (fast formation, well separated in space, stabilized to prevent back transfer) for the potential sensitivity of cryptochromes to weak magnetic fields. The interaction of the radical pair with the earth' magnetic field provides the basis for the hypothesis that cryptochromes are responsible for the magnetic compass sense of migratory birds [8].

Furthermore, for both *A. thal.* cryptochrome and *E. coli* photolyase [126] a structural "locking" mechanism was found that enforces the electron reaction to be one-way and enhances the stabilization already provided by the strong exothermicity of the reaction itself. This mechanism provides additional stabilization of RP-C on the nanosecond timescale in *E. coli* photolyase, while it allows stabilization of RP-B in *A. thal* cryptochrome, each time allowing to fine-tune the radical pair stabilization specific to the function of the individual protein. The "locking" mechanism was suggested to provide a qualitative answer to the different magnitudes of the magnetic field effect observed for the two proteins in experiment [68].

The current search for alternate electron transfer pathways in cryptochromes fueled the idea of refining the the parametrization of tryptophan and tyrosine residues. The DFTB energies deviate slightly, by about 0.2 eV, from the results obtained with higher-level methods which might affect the electron transfer dynamics. It was shown that the DFTB HOMO energies nicely reproduce the dynamical features of the ionization potential calculated with higher-level methods, and thus were modified by applying an additive shift. An improved description of the electron transfer energetics was obtained, which can be tested for example by applying the direct electron transfer approach to *X. laevis* cryptochrome where alternate paths were detected experimentally [163]. Furthermore, proposed alternate electron transfer pathways in *A. thal.* cryptochrome-2 [60] and the effect of small molecular activators, such as ATP, on their efficiency can be easily investigated with the direct electron transfer approach.

The simulation of the DNA photo-repair required the development of new model parameters. While the photo-activation reaction was modeled as hole transfer along HOMO orbitals, the simulation of DNA repair required the description of the excess electron as linear combination of LUMO orbitals. Comparison of the DFTB energies to the results of higher-level methods showed that the DFTB LUMO energies are well suited to describe the electron transfer energetics. A small additive shift was necessary to modify the DFTB energies. The results also suggested that the molecular dipole moment of the flavin and the thymine dimer changes several times during the photo-repair. It was concluded that the varying dipole moments will likely induce dynamic repolarization of the environment and might effect the electron transfer dynamics substantially. However, direct electron transfer simulations need to be carried out to investigate the molecular

mechanism in detail.

The developed model parameters can be used to study the dynamics of the photo-repair, accounting for the complete interaction of the excess electron and the environment. Comparison of the results with existing experimental data will allow validation of the model. Dynamic effects such as solvent polarization and environmental relaxation can be analyzed on an atomic level and might provide new insight in the nature of the electron transfer mechanism. Additional residues that might be involved in the transfer, such as the adenine moiety of FAD, can be easily included in the model description. The efficiency of the repair reaction which is currently debated, might be investigated by analyzing the backward electron transfer from CPD^{-} to the flavin. Furthermore, non-equilibrium conformations, representing the situation when the electron transfer to CPD has been completed, but before the bond splitting occurs, can be obtained. These conformations might be used as a starting structure for subsequent potential free energy calculation which might help to solve the present disaccord between the experimental and theoretical results regarding the rate of ring opening.

List of Figures

List of Tables

Bibliography

[1] Taylor, J. S. *Acc. Chem. Res.* **27**, 76–82 (1994). 1

[2] Weber, S. *Biochim. Biophys. Acta - Bioenergetics.* **1707**, 1–23 (2005). 1, 3

[3] Setlow, R. B. *Science* **153**, 379–386 (1966). 1, 77, 95

[4] Müller, M. and Carell, T. *Curr. Opin. Struct. Biol.* **19**, 277 – 285 (2009). 1

[5] Ahmad, M. and Cashmore, A. R. *Phys. Med. Biol.* **30**, 851–861 (1996). 1

[6] Chaves, I., Pokorny, R., Byrdin, M., Hoang, N., Ritz, T., Brettel, K., Essen, L.-O., van der Horst, G. T., Batschauer, A., and Ahmad, M. *Annu. Rev. Plant Biol.* **62**, 335–364 (2011). 1, 2, 6

[7] Lin, C. and Todo, T. *Genome Biol.* **6**, 220 (2005). 1, 6

[8] Ritz, T., Adem, S., and Schulten, K. *Biophys. J.* **78**, 707–718 (2000). 2, 7, 70, 97

[9] Griffin, E. A., Staknis, D., and Weitz, C. J. *Science* **286**, 768–771 (1999). 2

[10] Zhu, H., Sauman, I., Yuan, Q., Casselman, A., Emery-Le, M., Emery, P., and Reppert, S. M. *PLoS Biol.* **6**, e4 01 (2008). 2

[11] Brudler, R., Hitomi, K., Daiyasu, H., Toh, H., Kucho, K., Ishiura, M., Kanehisa, M., Roberts, V. A., Todo, T., Tainer, J. A., and Getzoff, E. D. *Mol. Cell* **11**, 59 – 67 (2003). 2

[12] Selby, C. P. and Sancar, A. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 17696–17700 (2006). 2

[13] Pokorny, R., Klar, T., Hennecke, U., Carell, T., Batschauer, A., and Essen, L.-O. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 21023–21027 (2008). 2

[14] Douki, T. and Cadet, J. *Biochemistry* **40**, 2495–2501 (2001). 3

[15] Park, H. W., Kim, S. T., Sancar, A., and Deisenhofer, J. *Science* **268**, 1866–1872 (1995). 3

[16] Tamada, T., Kitadokoro, K., Higuchi, Y., Inaka, K., Yasui, A., de Ruiter, P., Eker, A., and Miki, K. *Nat. Struct. Biol.* **4**, 887–891 (1997). 3, 4

[17] Mees, A., Klar, T., Gnau, P., Hennecke, U., Eker, A. P. M., Carell, T., and Essen, L.-O. *Science* **306**, 1789–1793 (2004). 3, 78

[18] Humphrey, W., Dalke, A., and Schulten, K. *J. Mol. Graph.* **14**, 33–38 (1996). 4, 60

[19] Payne, G., Heelis, P., Rohrs, B., and Sancar, A. *Biochemistry* **26**, 7121–71217 (1987). 3

[20] Kim, S.-T. and Sancar, A. *Photochem. Photobiol.* **57**, 895–904 (1993). 3

[21] MacFarlane, A. and Stanley, R. J. *Biochemistry* **42**, 8558–8568 (2003). 3

[22] Kao, Y.-T., Saxena, C., Wang, L., Sancar, A., and Zhong, D. *Proc. Natl. Acad. Sci. U. S. A.* **45**, 1612816132 (2005). 3

[23] Liu, Z., Tan, C., Guo, X., Kao, Y.-T., Li, J., Wang, L., Sancar, A., and Zhong, D. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 14831–14836 (2011). 3, 5, 77, 95

[24] Thiagarajan, V., Byrdin, M., Eker, A., Müller, P., and Brettel, K. *Proc. Natl. Acad. Sci. U. S. A.* **23**, 9402–9407 (2011). 5, 95

[25] Masson, F., Laino, T., Rothlisberger, U., and Hutter, J. *Phys. Chem. Chem. Phys.* **10**, 400–410 (2009). 5

[26] Dou, Y., Xiong, S., Wu, W., Yuan, S., and Tang, H. *J. Photochem. Photobiol. B* **101**, 31–36 (2010).

[27] Hassanali, A. A., Zhong, D., and Singer, S. J. *J. Phys. Chem. B* **115**, 3860–3871 (2011). 5, 77

[28] Brettel, K. and Byrdin, M. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E1462 (2012). 5

[29] Zhong, D., Sancar, A., and Stuchebrukhov, A. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E1463 (2012). 5

[30] Antony, J., Medvedev, D. M., and Stuchebrukhov, A. A. *J. Am. Chem. Soc.* **122**, 1057–1065 (2000). 5

[31] Medvedev, D. M. and Stuchebrukhov, A. A. *J. Theor. Biol.* **210**, 237 – 248 (2001). 5

[32] Prytkova, T. R., Beratan, D. N., and Skourtis, S. S. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 802–807 (2007). 5

[33] Acocella, A., Jones, G. A., and Zerbetto, F. *J. Phys. Chem. B* **114**, 4101–4106 (2010). 5

[34] Liu, Z., Guo, X., Tan, C., Li, J., Kao, Y., Wang, L., Sancar, A., and Zhong, D. *J. Am. Chem. Soc.* **134**, 8104–8114 (2012). 5

[35] Wang, H., Chen, X., and Fang, W. *Phys. Chem. Chem. Phys.* **16**, 25432–25441 (2014). 5, 92

[36] Kavakli, I. H. and Sancar, A. *Biochemistry* **43**, 15103–15110 (2004). 6

[37] Brettel, K. and Byrdin, M. *Curr. Op. Struct. Biol.* **20**, 693–701 (2010). 6, 46, 51

[38] Li, Y. F., Heelis, P. F., and Sancar, A. *Biochemistry* **30**, 6322–6329 (1991). 6

[39] Cheung, M. S., Daizadeh, I., Stuchebrukhov, A. A., and Heelis, P. F. *Biophys. J.* **76**, 1241–1249 (1999). 6

[40] Aubert, C., Vos, M. H., Mathis, P., Eker, A. P. M., and Brettel, K. *Nature* **405**, 586–590 (2000). 6, 31

[41] Byrdin, M., Eker, A. P. M., Vos, M. H., and Brettel, K. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 8676–8681 (2003). 6, 31

[42] Byrdin, M., Villette, S., Eker, A. P. M., and Brettel, K. *Biochemistry* **46**, 10072–10077 (2007).

[43] Lukacs, A., Eker, A. P. M., Byrdin, M., Villette, S., Pan, J., Brettel, K., and Vos, M. H. *J. Phys. Chem. B* **110**, 15654–15658 (2006). 6

[44] Lukacs, A., Eker, A. P. M., Byrdin, M., Brettel, K., and Vos, M. H. *J. Am. Chem. Soc.* **130**, 14394–14395 (2008). 6, 31, 95

[45] Byrdin, M., Lukacs, A., Thiagarajan, V., Eker, A. P. M., Brettel, K., and Vos, M. H. *J. Phys. Chem. A* **114**, 3207–3214 (2010). 6, 31, 32, 95

[46] Liu, Z., Tan, C., Guo, X., Li, J., Wang, L., Sancar, A., and Zhong, D. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 12966–12971 (2013). 6, 47, 48, 77, 96

[47] Liu, Z., Tan, C., Guo, X., Li, J., Wang, L., and Zhong, D. *Phys. Chem. Lett.* **5**, 820–825 (2014). 6, 47

[48] Yang, H.-Q., Wu, Y.-J., Tang, R.-H., Liu, D., Liu, Y., and Cashmore, A. R. *Cell* **103**, 815 – 827 (2000). 6

[49] Liu, Y., Li, X., Li, K., Liu, H., and Lin, C. *PLoS Genet.* **9**, e1003861 (2013). 6

[50] Cashmore, A. R. *Cell* **114**, 537 – 543 (2003). 6

[51] Liscum, E., Hodgson, D. W., and Campbell, T. J. *Plant Physiol.* **133**, 1429–1436 (2003). 6

[52] Müller, P. and Ahmad, M. *J. Biol. Chem.* **286**, 21033–21040 (2011). 6

[53] Burney, S., Wenzel, R., Kottke, T., Roussel, T., Hoang, N., Bouly, J., Bittl, R., Heberle, J., and Ahmad, M. *Angew. Chem., Int. Ed.* **51**, 93569360 (2012). 7

[54] Hoang, N., Schleicher, E., Kacprzak, S., Bouly, J., Picot, M., Wu, W., Berndt, A., Wolf, E., Bittl, R., and Ahmad, M. *PLoS Biol.* **6**, e160 (2008). 7

[55] Vaidya, A. T., Top, D., Manahan, C. C., Tokuda, J. M., Zhang, S., Pollack, L., Young, M. W., and Crane, B. R. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 20455–20460 (2013). 7

[56] Song, S.-H., Öztürk, N., Denaro, T. R., Arat, N. z., Kao, Y.-T., Zhu, H., Zhong, D., Reppert, S. M., and Sancar, A. *J. Biol. Chem.* **282**, 17608–17612 (2007). 7, 71

[57] Ozturk, N., Selby, C. P., Annayev, Y., Zhong, D., and Sancar, A. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 516–521 (2011). 7, 71

[58] Müller, P., Bouly, J. P., Hitomi, K., Balland, V Getzoff, E. D., Ritz, T., and Brettel, K. *Sci. Rep.* **4**, 5175 (2014). 7, 68, 69, 71, 96

[59] Caillez, F., Müller, P., Gallois, M., and de la Lande, A. *J. Am. Chem. Soc.* **136**, 12974–12986 (2014). 7, 68, 69

[60] Engelhard, C., Wang, X., Robles, D., Moldt, J., Essen, L.-O., Batschauer, A., Bittl, R., and Ahmad, M. *The Plant Cell Online* **26**, 4519–4531 (2014). 7, 71, 97

[61] Banerjee, R., Schleicher, E., Meier, S., Viana, R. M., Pokorny, R., Ahmad, M., Bittl, R., and Batschauer, A. *J. Biol. Chem.* **282**, 14916–14922 (2007). 7, 51

[62] Berndt, A., Kottke, T., Breitkreuz, H., Dvorsky, R., Hennig, S., Alexander, M., and Wolf, E. *J. Biol. Chem.* **282**, 13011–13021 (2007). 7, 51

[63] Kottke, T., Batschauer, A., Ahmad, M., and Heberle, J. *Biochemistry* **45**, 2472–2479 (2006). 7, 51

[64] Langenbacher, T., Immeln, D., Dick, B., and Kottke, T. *J. Am. Chem. Soc.* **131**, 14274–14280 (2009). 7, 51, 68, 96

[65] Giovani, B., Byrdin, M., Ahmad, M., and Brettel, K. *Nature Struct. Biol.* **10**, 489–490 (2003). 7

[66] Immeln, D., Weigel, A., Kottke, T., and J. Luis Pérez Lustres. *J. Am. Chem. Soc.* **134**(30), 12536–12546 (2012). 7, 54, 58, 59, 68, 69, 71, 95

[67] Solov'yov, I. A., Domratcheva, T., Moughal Shahi, A. R., and Schulten, K. *J. Am. Chem. Soc.* **134**, 18046–18052 (2012). 7, 53, 62, 66, 68, 96

[68] Maeda, K., Robinson, A. J., Henbest, K. B., Hogben, H. J., Biskup, T., Ahmad, M., Schleicher, E., Weber, S., Timmel, C. R., and Hore, P. J. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 4774–4779 (2012). 7, 58, 70, 97

[69] Rodgers, C. and Hore, P. J. *Proc. Natl. Acad. Sci. USA* **106**, 353–360 (2009). 7

[70] Liedvogel, M., Maeda, K., Henbest, K., Schleicher, E., Simon, T., Timmel, C. R., Hore, P., and Mouritsen, H. *PLoS ONE* **2**, e1106 (2007).

[71] Mouritsen, H. *Nature* **484**, 320–321 (2012).

[72] Dodson, C. A., Hore, P., and Wallace, M. I. *Trends Biochem. Sci.* **38**, 435–446 (2013). 7, 8, 51, 57

[73] Rodgers, C. *Pure Appl. Chem.* **81**, 19–43 (2009). 7

[74] Steiner, U. E. and Ulrich, T. *Chem. Rev.* **89**, 51–147 (1989). 7

[75] Schulten, K., Swenberg, C., and Weller, A. *Z. Phys. Chem. (Muenchen, Ger.)* **111**, 1–5 (1978). 7

[76] Wolfram Koch, M. C. H. *A Chemist's Guide to Density Functional Theory*. Wiley-VCH Verlag GmbH, (2001). 11, 15, 79, 80

[77] Szabo, A. and Ostlund, N. S. *Modern Quantum Chemistry*. Dover Publications, (1996).

[78] Leach, A. R. *Molecular Modelling. Principles and Application*. Addison Wesley Longman Limited, (1996). 11, 20, 21

[79] Marcus, R. A. *J. Chem. Phys.* **24**, 966–978 (1956). 11, 23, 95

[80] Marcus, R. A. *J. Chem. Phys.* **24**, 979–989 (1956). 11, 23, 95

[81] Kubař, T. and Elstner, M. *J. R. Soc. Interface* **10**, 20130415 (2013). 11, 44

[82] Kubař, T. and Elstner, M. *J. Phys. Chem. B* **114**, 11221–11240 (2010). 11, 25, 29, 81, 95

[83] Slater, J.; Verma, H. *Phys. Rev.* **34**, 12931295 (1929). 13

[84] Koopmans, T. C. *Physica* **1**, 104 (1934). 14

[85] Hohenberg, P. and Kohn, W. *Phys. Rev.* **136**, B864–B871 (1964). 14

[86] Kohn, W. and Sham, L. J. *Phys. Rev.* **140**, A1133–A1138 (1965). 15

[87] Perdew, J. P., Burke, K., and Ernzerhof, M. *Phys. Rev. Lett.* **77**, 3865–3868 (1996). 15, 17

[88] Becke, A. D. *J. Chem. Phys.* **98**, 5648–5652 (1993). 15

[89] Chai, J.-D. and Head-Gordon, M. *Phys. Chem. Chem. Phys.* **10**, 6615–6620 (2008). 15, 16, 79

[90] Janak, J. F. *Phys. Rev. B* **18**, 7165–7168 (1978). 15, 28

[91] Politzer, P.; Abu-Awwad, F. A. *Theoret. Chem. Acc.* **99**, 83–87 (1998). 15, 73

[92] Perdew, J. P., Parr, R. G., Levy, M., and Balduz, J. L. *Phys. Rev. Lett.* **49**, 1691–1694 (1982). 16

[93] Gaus, M., Cui, Q., and Elstner, M. *J. Chem. Theory Comput.* **7**, 931–948 (2011). 16, 17, 53, 66, 72, 79

[94] Porezag, D., Frauenheim, T., Köhler, T., Seifert, G., and Kaschner, R. *Phys. Rev. B* **51**(19), 12947–12957 May (1995). 16

[95] Elstner, M., Porezag, D., Jungnickel, G., Elsner, J., Haugk, M., Frauenheim, T., Suhai, S., and Seifert, G. *Phys. Rev. B* **58**, 7260 (1998). 17, 26

[96] Gaus, M. *J. Chem. Theory Comput.* **9**, 931–948 (2013). 17, 66, 67

[97] Koskinen, P. and Mäkinen, V. *Comp. Mater. Sci.* **47**, 237–253 (2009). 17

[98] Cornell, W., Cieplak, P., Bayly, C., Gould, I., Merz, K. J., Ferguson, D., Spellmeyer, D., Fox, T., Caldwell, J., and Kollman, P. *J. Am. Chem. Soc.* **117**, 51795197 (1995). 18

[99] Verlet, L. *Phys. Rev.* **159**, 98 (1967). 19

[100] Hockney, R. *Method. Comp. Phys.* **9**, 135–211 (1970). 19

[101] Evans, D. J. and Holian, B. L. *J. Chem. Phys.* **83**, 4069–4074 (1985). 20, 33, 53, 78

[102] Parinello, M. and Rahman, A. *J. Appl. Phys.* **52**, 7182–7190 (1981). 20, 33, 53, 78

[103] Hub, J. S., de Groot, B. L., and van der Spoel, D. *J. Chem. Theory Comput.* **6**, 3713–3720 (2010). 21, 54

[104] Skourtis, S. S. *Peptide Science* **100**, 82–92 (2013). 21, 22, 23

[105] Marcus, R. A. *Ann. Rev. Phys. Chem.* **15**, 155–196 (1964). 23, 95

[106] Marcus, R. A. and Sutin, N. *Biochim. Biophys. Acta* **811**, 265 – 322 (1985). 23, 95

[107] Levich, V. G. and Dogonadze, R. R. *Dokl. Akad. Nauk SSSR* **124**, 123–126 (1959). 23

[108] Hush, N. S. *Trans. Farad. Soc.* **57**, 557–580 (1961).

[109] Hopfield, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **71**, 3640–3644 (1974).

[110] Jortner, J. *J. Chem. Phys.* **64**, 4860–4867 (1976).

[111] Moser, C. C., Keske, J. M., Warncke, K., Farid, R. S., and Dutton, P. L. *Nature* **355**(6363), 796–802 (1992). 23

[112] Skourtis, S. S., Waldeck, D. H., and Beratan, D. N. *Annu. Rev. Phys. Chem.* **61**, 461–485 (2010). 25

[113] Beratan, D. N., Skourtis, S. S., Balabin, I. A., Balaeff, A., Keinan, S., Venkatramani, R., and Xiao, D. *Acc. Chem. Res.* **42**, 1669–1678 (2009). 25

[114] Balabin, I. A., Beratan, D. N., and Skourtis, S. *Phys. Rev. Lett.* **101**, 49–53 (2008).

[115] Grozema, F. C., Siebbeles, L. D. A., Berlin, Y. A., and Ratner, M. A. *Phys. Chem. Chem. Phys.* **3**, 536–539 (2002).

[116] Grozema, F., Berlin, Y., and Siebbeles, L. D. A. *J. Am. Chem. Soc.* **122**, 10903–10909 (2000). 25

[117] Hatcher, E., Balaeff, A., Keinan, S., Venkatramani, R., and Beratan, D. N. *J. Am. Chem. Soc.* **130**, 11752–11761 (2008). 25

[118] Wang, H., Lin, S., Allen, J. P., Williams, J. C., Blankert, S., Laser, C., and Woodbury, N. W. *Science* **316**, 747–750 (2007). 25

[119] Skourtis, S. S. and Beratan, D. N. *Science* **316**, 703–704 (2007). 25

[120] Ehrenfest, P. *Z. Phys.* **45**, 455–457 (1927). 25

[121] Tully, J. C. and Preston, R. K. *J. Chem. Phys.* **55**, 562–572 (1971). 25

[122] Tully, J. C. *J. Chem. Phys.* **93**, 1061–1071 (1990). 25

[123] Kubař, T., Woiczikowski, P. B., Cuniberti, G., and Elstner, M. *J. Phys. Chem. B* **112**, 7937–7947 (2008). 25, 71, 74

[124] Kubar, T. and Elstner, M. *Phys. Chem. Chem. Phys.* **15**, 5794–5813 (2013). 32, 95

[125] Kubar, T. and Elstner, M. *J. Royal Soc. Interf.* **10**, 20130415 (2013). 25, 95

[126] Woiczikowski, P. B., Steinbrecher, T., Kubař, T., and Elstner, M. *J. Phys. Chem. B* **115**, 9846–9863 (2011). 30, 31, 32, 33, 35, 38, 39, 40, 45, 48, 49, 53, 62, 63, 64, 65, 69, 77, 94, 96, 97

[127] Krapf, S., Koslowski, T., and Steinbrecher, T. *Phys. Chem. Chem. Phys.* **12**, 9516–9525 (2010). 31

[128] Furse, K. E. and Corcelli, S. A. *Phys. Chem. Lett.* **1**, 1813–1820 (2010). 32, 48

[129] Golosov, A. A. and M., K. *J. Phys. Chem. B* **111**, 1482–1490 (2007). 32, 48

[130] Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. *J. Comput. Chem.* **25**, 1157–1174 (2004). 32, 53, 78

[131] Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. *Proteins* **65**, 712–725 (2006). 32, 53, 78

[132] Meagher, K. L., Redman, L. T., and Carlson, H. A. *J. Comput. Chem.* **24**, 1016–1025 (2003). 32, 53, 78

[133] Bayly, C. I., Cieplak, P., Cornell, W. D., and Kollman, P. A. *J. Phys. Chem.* **97**, 10269 (1993). 32, 53, 78

[134] Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. *J. Comput. Chem.* **18**, 1463–1472 (1997). 33, 53, 78

[135] Aubert, C., Mathias, P., Eker, A., and Brettel, K. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 5423–5427 (1999). 33

[136] Kao, Y.-T., Tan, C., Song, S.-H., Öztürk, N., Li, J., Wang, L., Sancar, A., and Zhong, D. *J. Am. Chem. Soc.* **130**, 7695–7701 (2008). 35

[137] Blumberger, J. and Lamoureux, G. *Mol. Phys.* **106**, 1597–1611 (2008). 44

[138] Alden, R. G., Parson, W. W., Chu, C. T., and Warshel, A. *J. Am. Chem. Soc.* **117**, 12284–12298 (1995). 44

[139] Ceccarelli, M. and Marchi, M. *J. Phys. Chem. B* **107**, 5630–5641 (2003). 44

[140] Blumberger, J. *Phys. Chem. Chem. Phys.* **10**, 5651–5667 (2008). 44

[141] Schulten, K. and Tesch, M. *Chem. Phys.* **158**, 421–446 (1991). 44

[142] Leontyev, I. and Stuchebrukhov, A. *Phys. Chem. Chem. Phys.* **13**, 2613–2626 (2011). 48, 96

[143] Giovani, B., Byrdin, M., Ahmad, M., and Brettel, K. *Nat. Struct. Biol.* **10**, 489–490 (2003). 51

[144] Zeugner, A., Byrdin, M., Bouly, J.-P., Bakrim, N., Giovani, B., Brettel, K., and Ahmad, M. *J. Biol. Chem.* **280**, 19437–19440 (2005). 51, 56, 59

[145] Brautigam, C. A., Smith, B. S., Ma, Z., Palnitkar, M., Tomchick, D. R., Machius, M., and Deisenhofer, J. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 12142–12147 (2004). 52

[146] Lüdemann, G., Solov'ov, I. A., Kubař, T., and Elstner, M. *J. Am. Chem. Soc.* **137**, 1147–1156 (2015). 52, 77, 94, XXVII

[147] Solov'yov, I. A., Domratcheva, T., and Schulten, K. *Sci. Rep.* **4**, 3845 (2014). 51, 57, 62

[148] Song, Y., Mao, J., and Gunner, M. R. *J. Comput. Chem.* **30**, 2231–2247 (2009). 53

[149] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. *J. Chem. Phys.* **79**, 926 (1983). 53, 78

[150] Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. *J. Chem. Theory Comput.* **4**, 435–447 (2008). 53, 78

[151] Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. *J. Comput. Chem.* **26**, 1781–1802 (2005). 53

[152] MacKerell Jr., A. D., Bashford, D., Bellott, M., Dunbrack Jr., R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, I. W. E., Roux, B., Schlenkrich, M., Smith, J., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. *J. Phys. Chem. B* **102**, 3586–3616 (1998). 53, 68

[153] MacKerell Jr., A. D., Feig, M., and Brooks III, C. L. *J. Comput. Chem.* **25**, 1400–1415 (2004). 53, 68

[154] Elstner, M., Porezag, D., Jungnickel, G., Elsner, J., Haugk, M., Frauenheim, T., Suhai, S., and Seifert, G. *Phys. Rev. B* **58**, 7260–7268 (1998). 53, 72, 79, 96

[155] Solov'yov, I. A., Hore, P. J., Ritz, T., and Schulten, K. *Quantum Effects in Biology*, chapter A chemical compass for bird navigation, 218–236. Cambridge University Press, Cambridge (2014). 56

[156] Efimova, O. and Hore, P. *Biophys. J.* **94**, 1565–1574 (2008). 57

[157] Sancar, A. *Chem. Rev.* **103**, 2203–2237 (2003). 63

[158] Lüdemann, G., Woiczikowski, P. B., Kubař, T., Elstner, M., and Steinbrecher, T. B. *J. Phys. Chem. B* **117**, 10769–10778 (2013). 63, 69, 71, 77, 94

[159] Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, Jr., J. A., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., 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, J. E., Hratchian, H. P., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P., Dannenberg, J. J., Zakrzewski, V. G., Dapprich, S., Daniels, A. D., Strain, M. C., Farkas, O., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Ortiz, J. V., Cui, Q., Baboul, A. G., Clifford, S., Cioslowski, J., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Gonzalez, C., and Pople, J. A. (2003). Gaussian, Inc., Wallingford, CT, 2004. 67, 72

[160] Herzfeld, K. F. *Ann. Phys.* **59**, 635–667 (1919). 68

[161] Lee, A. A., Lau, J. C., Hogben, H. J., Biskup, T., Kattnig, D. R., and Hore, P. J. *J. R. Soc. Interface* **11**, 20131063 (2014). 70

[162] Li, X., Wang, Q., Yu, X., Liu, H., Yang, H., Zhao, C., Liu, X., Tan, C., Klejnot, J., Zhong, D., and Lin, C. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 20844–20849 (2011). 71

[163] Biskup, T., Paulus, B., Okafuji, A., Hitomi, K., Getzoff, E. D., Weber, S., and Schleicher, E. *J. Biol. Chem.* **288**, 9249–9260 (2013). 71, 97

[164] Aradi, B., Hourahine, B., and Frauenheim, T. *J. Phys. Chem. A* **111**, 5678–5684 (2007). 72, 79

[165] Rasmussen, E., (2013). private communications. 72

[166] Troisi, A. and Orlandi, G. *J. Phys. Chem. B* **106**, 2093–2101 (2002). 74

[167] Voityuk, A. A. *J. Chem. Phys.* **128**(11), 115101 (2008). 74

[168] Saettel, N. J. and Wiest, O. *J. Am. Chem. Soc.* **123**, 2693–2694 (2001). 77

[169] Masson, F., Laino, T., Tavernelli, I., Rothlisberger, U., and Hutter, J. *J. Am. Chem. Soc.* **130**, 3443–3450 (2008). 77

[170] Kneuttinger, A. C., Kashiwazaki, G., Prill, S., Heil, K., M"uller, M., and Carell, T. *Photochem. Photobiol.* **90**, 1–14 (2014). 77

[171] Lee, J.-H., Park, C.-J., Shin, J.-S., Ikegami, T., Akutsu, H., and Choi, B.-S. *Nucleic Acids Res.* **32**, 2474–2481 (2004). 78

[172] Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, Jr., J. A., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, O., Foresman, J. B., Ortiz, J. V., Cioslowski, J., and Fox, D. J. Gaussian Inc. Wallingford CT 2009. 79, 91

[173] McLean, A. D. and Chandler, G. S. *J. Chem. Phys.* **72**, 5639–48 (1980). 79

[174] Kendall, R. A., Jr., T. H. D., and Harrison, R. J. *J. Chem. Phys.* **96**, 6796–806 (1992). 79

[175] Zhan, C.-G., Nichols, J. A., and Dixon, D. A. *J. Phys. Chem. A* **107**, 4184–4195 (2003). 80, 81

[176] Roesch, N. and Trickey, S. *J. Chem. Phys.* **106**, 8940 (1997). 80

[177] Rienstra-Kiracofe, J. C., Tschumper, G. S., Schaefer, H. F., Nandi, S., and Ellison, G. B. *Chem. Rev.* **102**, 231–282 (2002). 80

[178] Arduengo, A. J., Bock, H., Chen, H., Denk, M., Dixon, D. A., Green, J. C., Herrmann, W. A., Jones, N. L., Wagner, M., and West, R. *J. Am. Chem. Soc.* **116**, 6641–6649 (1994). 81

[179] Heelis, P. F., Hartman, R. F., and Rose, S. D. *Chem. Soc. Rev.* **24**, 289–297 (1995). 82

[180] Sancar, A. *Biochemistry* **33**, 2–9 (1994). 82

[181] Barbatti, M. *Phys. Chem. Chem. Phys.* **15**, 3342–3354 (2014). 83, 89, 91

[182] Desfrancois, C., Abdoul-Carime, H., Carles, S., Periquet, V., Schermann, J. P., Smith, D. M. A., and Adamowicz, L. *J. Chem. Phys.* **110**, 11876–11883 (1999). 83, 93

[183] Crawford, O. H. and Garrett, W. R. *J. Chem. Phys.* **66**, 4968–4970 (1977). 83

[184] Schiedt, J., Weinkauf, R., Neumark, D., and Schlag, E. *Chem. Phys.* **239**, 511 – 524 (1998). 83

[185] TURBOMOLE V6.2 2010, a development of University of Karlsruhe, Forschungszentrum Karlsruhe GmbH, 1989-2007, T., (since 2007). available from . 91

[186] F. Weigend, F. F. and J., R. A. *J. Chem. Phys.* **119**, 12753 (2003). 91

[187] Klamt, A. and Schuurmann, G. *J. Chem. Soc., Perkin Trans. 2* **5**, 799–805 (1993). 91

[188] Chang, C.-W., Guo, L., Kao, Y.-T., Li, J., Tan, C., Li, T., Saxena, C., Liu, Z., Wang, L., Sancar, A., and Zhong, D. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 2914–2919 (2010). 92

[189] Marazzi, M., (2015). private communications. 92, XVIII

[190] Henbest, K. B., Maeda, K., Hore, P. J., Joshi, M., Bacher, A., Bittl, R., Weber, S., Timmel, C. R., and Schleicher, E. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14395–14399 (2008). 95

Appendix

A RMSD of the Cryptochrome Resting State

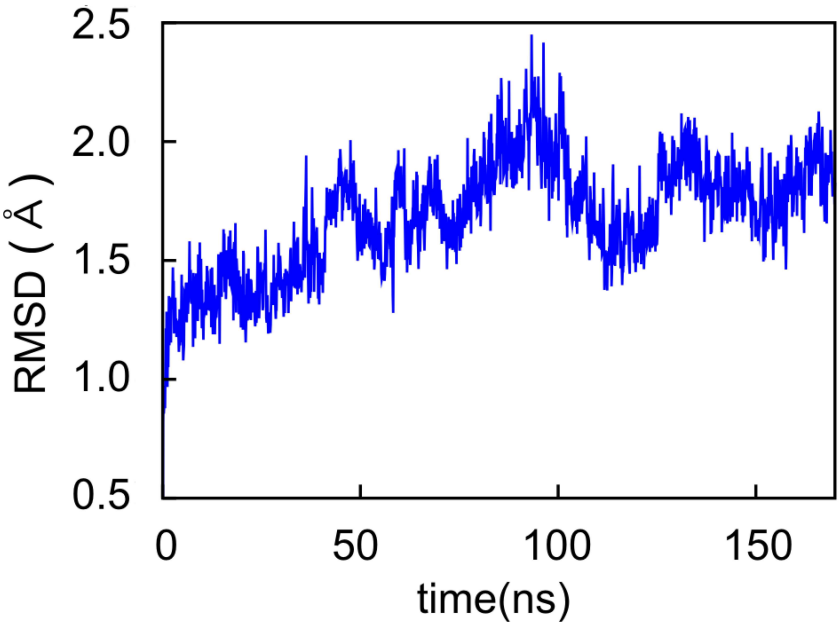


Figure .1: Time evolution of the cryptochrome RMSD. *The values of the RMSD are computed for the protein backbone atoms in reference to the protein resting state.*

B RMSD of Photolyase in Complex with DNA

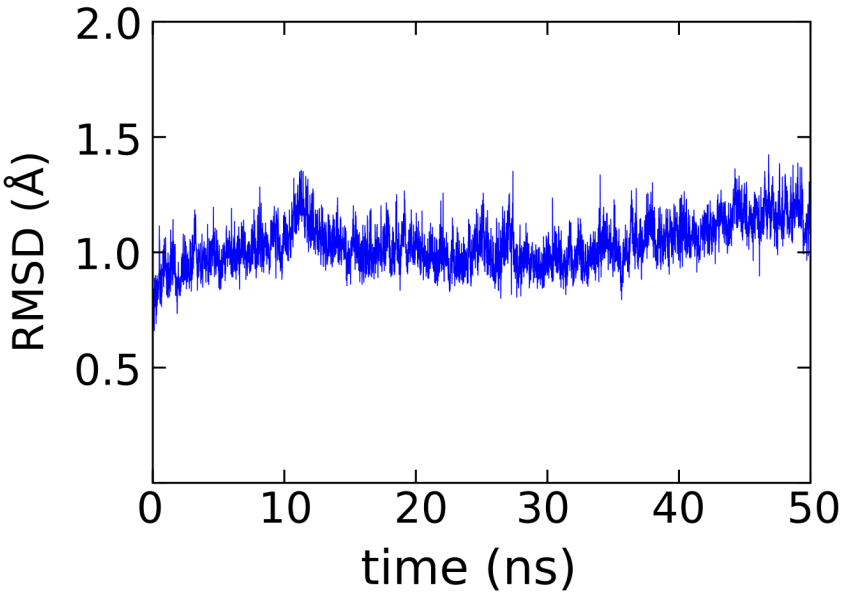


Figure .2: Time evolution of the RMSD of photolyase in complex with DNA. *The values of the RMSD are computed for the protein backbone atoms in reference to the protein inactive state.*

C Quantum Chemical Characterization of the Flavin Excited State

The vertical singlet excitation energies of $FADH^-$ have been determined at different levels of theory and are represented in Tab. reftab:excited. The energies computed for the first singlet excited state are in good agreement with the experimental data. The vertical singlet excitation energy of the first excited state calculated with TDDFT is overestimated by about 0.3 eV compared to the experimental value, CASPT2 and SORCI are in slightly better agreement with the experiment, deviating by about 0.1 eV and 0.2 eV from the experimental result.

Table .1: *Excited state energies in vaccum of $FADH^-$ computed with various methods. Energies are given in eV.*

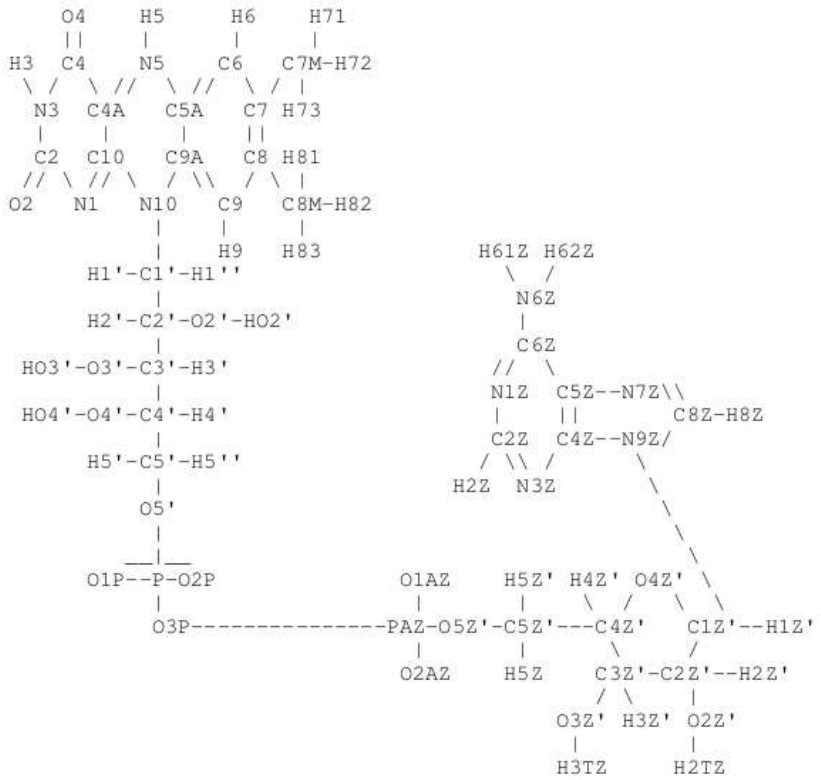
electronic state	TDDFT(B3LYP-631G*)	CASPT2 [189]	SORCI (HF/def2-TZVP)
S1	2.86	2.58	2.63
S2	3.32	3.08	3.13
S3	4.37	4.78	5.38

D Dipole Moments of Flavin and Thymine Dimer

Table .2: *Dipole moments of flavin and thymin dimer in different states calculated with DFT and HF methods. Energies are given in eV.*

	DFTB	PBE	B3LYP	HF	ω97XD/cc-pVTZ	Exp.
$FADH^-$	-	10.07	10.78	12.33	11.35	
$FADH^\bullet$	8.45	9.11	9.53	9.49	9.52	
CPD	6.27	6.12	6.50	7.54	6.67	
$CPD^{\bullet-}$	-	6.62	7.08	8.43	7.48	

E Atomic Charges



		FAD	FAD^-	$FADH^\bullet$	$FADH^-$
ATOM	N1	-0.868	-0.680	-0.7186	-0.6212
ATOM	C2	1.067	0.895	0.8913	0.8317
ATOM	O2	-0.649	-0.744	-0.6118	-0.7010
ATOM	N3	-0.849	-0.704	-0.7130	-0.6901
ATOM	H3	0.407	0.357	0.3878	0.3628
ATOM	C4	0.781	0.669	0.7010	0.6811
ATOM	O4	-0.578	-0.660	-0.5829	-0.6872
ATOM	C4A	0.127	0.035	-0.0907	-0.1141
ATOM	N5	-0.518	-0.623	-0.3907	-0.7141
ATOM	H5	-	-	0.3956	0.3587
ATOM	C5A	0.457	0.568	0.2029	0.4676
ATOM	C6	-0.392	-0.449	-0.4186	-0.5604
ATOM	H6	0.219	0.186	0.2226	0.2170
ATOM	C7	0.119	0.162	0.1606	0.1799
ATOM	C7M	-0.287	-0.290	-0.3234	-0.2375
ATOM	H71	0.094	0.077	0.1006	0.0664
ATOM	H72	0.094	0.077	0.1006	0.0664
ATOM	H73	0.094	0.077	0.1006	0.0664
ATOM	C8	0.174	0.084	0.1386	0.0337
ATOM	C8M	-0.345	-0.287	-0.3228	-0.2510
ATOM	H81	0.109	0.077	0.1021	0.0673
ATOM	H82	0.109	0.077	0.1021	0.0673
ATOM	H83	0.109	0.077	0.1021	0.0673
ATOM	C9	-0.381	-0.346	-0.3623	-0.3788
ATOM	H9	0.274	0.237	0.2679	0.2488
ATOM	C9A	-0.065	-0.346	0.0779	-0.0549
ATOM	N10	0.021	0.092	-0.1200	0.0511
ATOM	C10	0.551	0.188	0.5406	0.2345
ATOM	C1’	-0.246	-0.146	-0.1421	-0.0733
ATOM	H1’	0.149	0.089	0.1134	0.0620
ATOM	H1”	0.149	0.089	0.1134	0.0620
ATOM	C2’	0.144	0.368	0.1898	-0.0658
ATOM	H2’	-0.089	-0.004	0.0770	0.1387
ATOM	O2’	-0.604	-0.760	-0.7009	-0.6203
ATOM	HO2’	0.404	0.450	0.4430	0.3710
ATOM	C3’	0.277	0.384	0.2087	0.3583
ATOM	H3’	0.013	0.021	0.0722	0.0050
ATOM	O3’	-0.659	-0.708	-0.7611	-0.7110
ATOM	HO3’	0.439	0.424	0.4968	0.4189
ATOM	C4’	-0.112	0.089	0.3383	0.4710
ATOM	H4’	0.014	0.009	0.0417	-0.0638
ATOM	O4’	-0.646	-0.721	-0.7802	-0.7490
ATOM	HO4’	0.447	0.454	0.4641	0.4125
ATOM	C5’	-0.063	-0.045	-0.3804	-0.3180
ATOM	H5’	-0.051	0.037	0.1011	0.0661
ATOM	H5”	-0.051	0.037	0.1011	0.0661
ATOM	O5’	-0.125	-0.125	-0.125	-0.125

		ADP
ATOM	P	1.385
ATOM	O1P	-0.889
ATOM	O2P	-0.889
ATOM	O3P	-0.569
ATOM	C4Z'	0.107
ATOM	H4Z'	0.117
ATOM	O4Z'	-0.355
ATOM	C1Z'	0.039
ATOM	H1Z'	0.201
ATOM	C5Z	0.052
ATOM	N7Z	-0.607
ATOM	C8Z	0.201
ATOM	H8Z	0.155
ATOM	N9Z	-0.025
ATOM	N1Z	-0.761
ATOM	C2Z	0.588
ATOM	H2Z	0.047
ATOM	N3Z	-0.700
ATOM	C4Z	0.305
ATOM	C6Z	0.701
ATOM	N6Z	-0.902
ATOM	H61Z	0.412
ATOM	H62Z	0.412
ATOM	C2Z'	0.067
ATOM	H2Z'	0.097
ATOM	O2Z'	-0.614
ATOM	H2TZ	0.419
ATOM	C3Z'	0.202
ATOM	H3Z'	0.062
ATOM	O3Z'	-0.654
ATOM	H3TZ	0.438
ATOM	C5Z'	0.056
ATOM	H5Z'	0.068
ATOM	H5Z	0.068
ATOM	O5Z'	-0.599
ATOM	PAZ	1.253
ATOM	O1AZ	-0.880
ATOM	O2AZ	-0.880

Copyright

Abbreviations

ADP	adenine phosphate
ATP	adenosine triphosphate
BOA	Born- Oppenheimer approximation
CPD	cyclobutane pyrimide dimer
DFT	density functional theory
DFTB	density functional tight binding
DNA	deoxyribonucleic acid
RMSD	root mean square fluctuation
EA	electron affinity
FAD	flavin adenine dinucleotide
GROMACS	Groningen Machine for Chemical Simulation
HOMO	highest occupied molecular orbital
IP	Ionization potential
KS	Kohn-Sham
LUMO	lowest unoccupied molecular orbital
MD	molecular dynamics
MM	molecular mechanics
PMF	potential of mean force
TD-DFT	time dependent density functional theory
QM/MM	quantum mechanics/molecular mechanics

Acknowledgement

First of all I would like to thank Prof. Marcus Elstner for giving me the opportunity to work on this interesting topic and for the helpful and continuous supervision during the last years. He always had time for constructive discussion and offered valuable scientifc advice while giving me the freedom to expand my research activities according to my own interest.

I am also very grateful to Dr. Tomás Kubař for continuous scientific support and interesting and helpful discussions. Specials thanks also go to Thomas Steinbrecher who partly supervised my work during his time as post-doc in Karlsruhe.

I would like to express my gratitude to Prof. Ilya A. Solov'yov (University of Southern Demark, Quantum Biology and Nanophysics Research Group) for the fruitful and on-going collaboration on the cryptochrome project [146].

Furthermore, I would like to thank my colleagues for the nice atmosphere in the group and especially Alexander Heck and Maximilian Kubillus for interesting science and non-science related discussions during coffee breaks. I am also thankful to Sabine Holthoff for helping with all administrative issues.

Many thanks also go to Tatiana Domratcheva (Max Planck Institute for Medical Research, Heidelberg) for helpful advice and discussion regarding the DNA repair project and Marco Marazzi (Université de Lorraine, Nancy,France) for providing the CASPT2 calculations of the flavin excited state.

Last but not least I would like to thank my boyfriend and my family for their continuous support.

Erklärung

Hiermit erkläre ich die vorliegende Arbeit selbständig angefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht und die Satzung der Universität Karlsruhe (TH) zur Sicherung guter wissenschaftlicher Praxis beachtet habe.

Karlsruhe, 01.06.2015

Gesa Lüdemann