Theoretical Studies on Photophysics and Photochemistry of DNA

Yuejie Ai

Doctoral Thesis in Theoretical Chemistry and Biology

School of Biotechnology

Royal Institute of Technology

Stockholm, Sweden 2011
To My Parents and Kerry
Abstract

Theoretical studies on biological systems like nucleic acid and protein have been widely developed in the past 50 years and will continue to be a topic of interest in frontiers of natural science. In addition to experimental science, computational modeling can give useful information and help us to understand biochemical issues at molecular, atomic and even electronic levels.

Deoxyribonucleic acid (DNA), the hereditary basis of life’s genetic identity, has always been a major topic of discussions since its structure was built in 1953. However, harmful UV radiation from sunlight can make damage to DNA molecules and eventually give rise to DNA damaging biological consequences, like mutagenesis, carcinogenesis, and cell death. Photostability, photodamage, and photorepair are of vital importance in the photophysics and photochemistry of DNA. In this thesis, we have applied high level computer-aided theoretical methods to explore the underlying mechanisms for these three critical issues of DNA. Special attentions are paid to the following aspects: the properties of the excited states, the design of relevant computational models and the effects of biological environments.

We have systematically studied the excited state properties of DNA from single base to base pair and oligonucleotides, where the concerted base pairing and base stacking effects were found to play important roles in DNA photostability. The UV-light induced isomerization mechanism between two photoproducts of DNA photodamage has been revealed in different biological environments. In association with DNA photodamage, the related photorepair processes have been proposed for different lesions in photolyase which is a catalytic enzyme for DNA, and the calculated results well explained the experimental observations. In particular, the internal and external properties of flavin cofactors have been extensively studied by combining the electronic structure and spectroscopic calculations. We have examined the effects of the intramolecular hydrogen bond on spectroscopic properties of flavins. The good agreements with the experimental spectra indicated that the biological self-regulation acted critical role in these biological systems.
Preface

The work presented in this thesis has been carried out at the Department of Theoretical Chemistry and Biology, School of Biotechnology, Royal Institute of Technology, Stockholm, Sweden.

Papers included in the thesis


**Paper VII. Yuejie Ai**, Guangjun Tian, Rongzhen liao, Qiong Zhang, Weihai Fang, and Yi Luo. Intrinsic natural property of flavin mononucleotide controls its optical spectra in three redox states. Manuscript.
List of Papers not Included in this Thesis


Comments on My Contribution to the Papers Included

- I have taken the major responsibility for the calculations and writing of paper I, III, IV, V, VI, VII.
- I was responsible for part of reaction pathway calculations of paper II.
Acknowledgements

It is an honor for me to give my acknowledgements here to all the people who provided their unselfish assistance.

First and foremost, I owe my deepest gratitude to my supervisor Prof. Yi Luo, whose guidance and kind encouragements enabled me to complete this thesis. I have benefited a great deal from not only his valuable scientific insights but also his positive views on life. With his support, I am much more confident of my academic research and also my future life. Also lots of thanks to his family: Dr. Kezhao Xing, dear Linda and Oscar. I cherish all the happy time we have spent together.

My special thanks go to Prof. Wei-Hai Fang, my former supervisor in Beijing Normal University. Thank you for giving me the opportunity to study in KTH. He is always my respectful teacher who firstly takes me to the fantastic theoretical world.

I would like to give my heartily thanks to Prof. Hans Ågren, the head of the department of theoretical chemistry and biology, for the kind consideration of our academic environment and requirements. Thanks for the lessons and help from Dr. Ying Fu, Prof. Fahmi Himo, Prof. Kersti Hermansson, Prof. Faris Gel'mukhanov, Dr. Yaoquan Tu, Dr. Olav Vahrtras, Dr. Zilvinas Rinkevicius.

This thesis would not have been possible without the shared discussion and collaboration with my dear classmates. They are Feng Zhang, Rongzhen Liao, Shilv Chen, Guangjun Tian. I am grateful for their unselfish work and shared assistance.

I would like to show my great thanks to my colleagues and friends: Qiong Zhang, Feng Zhang and Xing Chen, Igor Ying Zhang, Sai Duan, Lili Lin, Yuping Sun, Xiao Chen, Kai Fu, Guangjun Tian, Weijie Hua, Xin Li(both two), Zhijun Ning, Qiang Fu, Xiuneng Song, Yongfei Ji, Qiu Fang, .... The friendship between us is my most precious experience in Sweden.

Lastly, I offer my thanks and blessings to beloved parents and little brother, for their endless love and support over these years, and also my dear Kerry, for his selfless love.
Contents

Abstract......................................................................................................................I

Chapter 1 Introduction..........................................................................................1

Chapter 2 Photochemistry
  2.1 Photophysical and Photochemical Processes.............................................4
  2.2 Popular Computational Methods for Excited states.................................8

Chapter 3 Deoxyribonucleic acid (DNA) and its photochemistry
  3.1 The Building Blocks of DNA.................................................................16
  3.2 DNA Double Helix......................................................................................17
  3.3 DNA Photostability.....................................................................................19
  3.4 DNA Photodamage.....................................................................................25
  3.5 DNA Photorepair.........................................................................................28

Chapter 4 Summary of Papers
  4.1 Design of Theoretical Models for Biological Systems.........................40
  4.2 Computational Photophysical and Photochemical Studies on UV-induced DNA issues.................................................................51

References..............................................................................................................61
Chapter 1

Introduction

Light is one of the most important elements to life. The exploring of light has promoted the source of life perception since ancient times.\(^1\) For instance, the searching for the age-old question why plants can grow in the light finally leads to the discovery of the famous photosynthesis effect of plant, which is the most important photo-induced chemical reaction that involves the energy conversion between sunlight and human being.\(^1\) The photosynthesis effect is the early exploration for the photochemistry. Since then, scientists began to study extensive photochemical and photophysical changes when different substances interact with light and finally enriched the field of photochemistry. The development of photochemistry is closely related to two aspects: spectroscopic technology and related photochemical theories. From the flash photolysis technique\(^2\) to crossed molecular beam technique\(^5\) and femtosecond laser pulse technology\(^7\), the advanced experimental techniques provide a basis for the development of the photochemical experimental studies. At the same time, the theorists established first and second laws of photochemistry\(^9\), Lamber-Beer’s law\(^9\), Frank-Condon principle\(^10\), Jablonski diagram\(^11\), Fermi’s Golden rule\(^9\) that laid a firm foundation for understanding the complicated photophysical and photochemical processes. However, photochemical studies had been largely limited for small organic molecules for many years. The emergence of novel ultrafast femtosecond spectroscopy and supercomputer marks beginning of new era of the photochemistry, which has extended into many new areas, including artificial photosynthesis, visualization, light cycle (clock), solar energy, biosensor, and bio-imaging, just to name a few. The successful application of photochemistry to biology was highlighted by the Nobel Prize in chemistry 2008, which was awarded jointly to three scientists: Osamu Shimomura, Martin Chalfie and Roger Y. Tsien for
their outstanding contributions in Green Fluorescent Protein (GFP). An important branch of biological photochemistry, the nucleic acid photochemistry has always drawn much attention; Actually, the interaction between light and DNA is a thematic issue as old as time. Nevertheless, the photo science of DNA is inactive until the double-helical structure of DNA was discovered in 1953.\(^1\) Especially in recent years, the development of spectroscopic and computational techniques have brought immense interests into this field.\(^13\)-\(^15\)

When exposed to ultraviolet (UV)-light, the nucleic acids are promoted to the excited electronic states. Fortunately, DNA is intrinsically photo-stable and it can dissipate the excess electronic energy before photo-reaction happens.\(^15\) It is one kind of self-protection that keeps the maintenance of life. This photostability arises from the intrinsic geometric and electronic structures, which are the result of evolution. DNA is a double helix composed of base, pentose and phosphoric acid. Both the structural units and its specific space structure may be responsible for the photostability of DNA. There are great efforts by different groups in understanding the relaxation pathway of excited electronic states in DNA, and lots of proposed factors, such as base pairing, base stacking, exciton and excimers, or charge and proton transfer, are still under debate.\(^14\)-\(^15\) A much clearer picture of the underlying mechanism for DNA photostability is needed. On the other hand, although we are lucky most of the time, external hazard like UV radiation from sunlight or chemicals can sometimes make damage to DNA's integrity and eventually give rise to DNA mutations and cancer risk.\(^16\) For instance, the high risk for skin cancer which is caused by the photodamage of skin cells’ DNA from UV radiation has attracted much more public attention nowadays.\(^17\) As estimated by American Cancer Society, skin cancer accounts for nearly half of all cancers in the United States. Photodamage of DNA is another practical topic in DNA photochemistry. Photochemical reactions in the excited states can produce harmful photoproducts. The absorption of the UV light by DNA results in mainly three photochemical lesions: cyclobutane pyrimidine dimers (CPDs), pyrimidine (6-4) pyrimidone [(6-4)PP] and its valence photoisomer, Dewar PP.\(^18\)-\(^24\) Compared to the extensively studied CPDs, the other two photoproducts have
been less studied. In response to UV photo-damage, cells can repair the photo-lesions via the photorepair process by photolyase.\textsuperscript{13} As the repair process is assisted by specific photolyase and the catalytic cofactors (flavins), things become much more complicated since both internal and external environmental factors are involved. It is a challenging task to explore the related static and dynamic features in those processes.

To identify new therapeutic strategies for the treatment of cancer arising from DNA photodamage and for the anti-cancer drug design, we need to know clearly the structures and properties of these harmful photolesions. More importantly, intensive studies on the formation mechanism of the DNA photolesions and DNA repair can prevent cancers and enhance the quality of our life. However, the underlying mechanisms of these behaviors are not always well understood. As a matter of common practice and the complexity of DNA systems, the current available experimental technology on DNA photochemistry has practical difficulties. Therefore, computer-aided theoretical studies can come in to play an important role in the practical application. In conjunction with experiments, we have employed high level computational methods such as the complete active-space self-consistent field (CASSCF)\textsuperscript{25-26}, complete active space with second-order perturbation theory (CASPT2)\textsuperscript{27}, CASPT2//CASSCF/Amber (QM/MM), time-dependent density functional theory (TDDFT)\textsuperscript{28-30} to investigate the structural properties and the possible mechanisms of the DNA photochemistry. For several systems, the actual profiles of optical spectra have been simulated with the inclusion of the vibronic coupling by using the linear coupling model (LCM).\textsuperscript{31}

The thesis is organized in the following way: In chapter 2, we will give a brief introduction to the basic photochemical concepts and the widely used computational methods for the excited states of photo-induced molecules. Research background and contents for the UV-induced DNA photochemistry are presented in chapter 3. Finally, a summary of the published papers are then given in chapter 4.
Chapter 2

Photochemistry

2.1 Photophysical and Photochemical Processes

As an important branch of chemistry, photochemistry mainly studies the photo-induced chemical phenomenon in molecules, atoms, and atomic ions, and so forth when they absorb and emit the energy of photons.\textsuperscript{1,9,32-36} Usually, photochemical processes are accompanied by photophysical processes. We have summarized those processes in the Jablonski diagram, see Figure 2.1. A Jablonski diagram is usually used to illustrate the varying molecular electronic states, energy levels, possible photophysical and photochemical processes.\textsuperscript{11,33-36}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{ Jablonski_Diagram.png}
\caption{Jablonski Diagram}
\end{figure}
In Jablonski diagram, the thick horizontal lines represent the electronic energy levels, while the thin and short lines are vibrational energy levels. In most cases, according to Pauli exclusion principle, electrons in the electronic ground state are paired (+1/2, -1/2). That is to say, S=0, M=2S+1=1. Most of the electronic ground states of molecules are singlet states, represented as $S_0$. With the absorption of photons, one of the paired electrons in the outmost layer is excited. According to the different spin directions of electrons, we classify them into different electronic excited states. If the electron keep the spin orientation along the excitation, M=2S+1=1, the excited state is still singlet state—"S" state. We use symbols $S_1$, $S_2$ to represent the first excited singlet state and the second excited singlet state, which are ordered by the magnitude of energies. On the other hand, the spin multiplicity M equals to 3 when they are parallel spins. In this case, the excited states are triplet states (T), represented as $T_1$, $T_2$….and so on. They are grouped according to multiplicity into horizontally displaced columns.

Usually, molecules can be excited to their electronic excited states by absorbing energy from ultraviolet light or visible light. In the electronic excited states, molecules are very active. Chemical reactions (CR) may occur if the excess energy can’t be released through radiative transitions, nonradiative transitions, and intermolecular energy transfer (ET).\(^{36}\)

Radiative transitions are transitions involve the absorption (A) and emission (F and P). They are displayed by straight solid arrows. The photons carry the different energy between the involved energy levels. Transition from the lowest excited electronic state is associated with emission of photons (light). If the transition occurs between states in the same spin, like from lowest singlet excited state ($S_1$) to the ground state ($S_0$), fluorescence may occur:

$$S_1 \rightarrow S_0 + h\nu (fluorescence)$$

The fluorescence is short-lived with the lifetime of $10^{-9}$-$10^{-6}$s. The radiative transition between different spin states, for example, transition from triplet $T_1$ state to $S_0$ state, can give rise to phosphorescence:
2.1 Photophysical and Photochemical Processes

\[ T_1 \rightarrow S_0 + h\nu(\text{phosphorescence}) \]

The lifetime of phosphorescence is \(10^{-4}-10^{-2}\) s, which is much longer than that of the fluorescence.

Nonradiative transitions are transitions not involved light emission. The excess energy is then dissipated in three main processes: vibrational relaxation, internal conversion (IC) in same spin states, and intersystem crossing (ISC) with different spin multiplicities. In vibrational relaxation, the molecule in a higher excited vibrational states returns to a lower vibrational state through energy transfer to the surroundings caused by collisions. IC and ISC are all radiationless transitions. It is worth to mention that the nonradiative transitions are essential for the deactivation process of the excited states, unless their rates are negligible compared with the radiative ones.

These are effective photophysical channels by which the excited molecule can return to the ground state. In competition with those processes, the photochemical reactions in terms of cleavage and formation of chemical bonds can occur in the excited state directly or in the ground state via nonradiative transitions.

To describe the photophysical and photochemical processes qualitatively, we need to determine the equilibrium structures and energies of the electronic excited states. The information of excited state is important to understand the properties of spectroscopy and other photochemical phenomena. The geometric relaxation and changes upon electronic excitation may reflect on the shape of spectra, fine structure, etc.. Energies and other molecular properties of excited states can be calculated for the analysis of spectroscopy and reaction mechanisms.

We take naphthalene for example in Figure 2.2. The lowest singlet and triplet excited states of naphthalene are \(S_1 (^1\pi\pi^*)\) and \(T_1 (^1\pi\pi^*)\) shown in Figure 2.2. After singlet-singlet absorption \((A)\) from ground state \(S_0\) to the \(S_1\) state, the first excited singlet state \((S_1)\) of naphthalene may decay to the \(T_1\) state through the intersystem crossing (ISC) or undergo the fluorescence \((F)\) to the ground state. The rate constant for the radiationless ISC process is \(6.6\times10^{-6}\) s\(^{-1}\) and it is comparable with the rate constant of fluorescence \((2\times10^{-6}\) s\(^{-1}\)).
Once the molecule is in the $T_1$ state, there are two ways for it to release the excess energy and return to the $S_0$ state. By spontaneously emitting a photon, the molecule can decay to the ground state by phosphorescence. The rate constant for this process is $0.02 \text{s}^{-1}$. Another intersystem crossing from $T_1$ to $S_0$ is much faster ($k=0.4 \text{ s}^{-1}$) than the phosphorescence. The fluorescence and the green phosphorescence can be readily observed by experimental technology. There is energy transfer process in the $T_1$ state involves a collision and electron exchange. In the excited states ($S_1$ and $T_1$ states), a series of photoreduction, photoaddition, and other photochemical reactions (CR) may take place. For example, the photochemical displacement of hydride ion by cyanide can take place after irradiation of naphthalene.

**Figure 2.2** Jablonski diagram for naphthalene, data from Ref.33
2.2 Popular Computational Methods for Excited States

Most of the photo-induced chemical reactions involve the electronic excited states of the molecules. Generally, there are two categories of quantum chemical approaches for calculation of excited states.\(^{39-42}\) First kind is the wave-function-based method, such as Configuration Interaction (CI) methods (Configuration Interaction with Single excitation (CIS), with Single and Double excitations (CISD), with Single, Double and Triplet excitations (CISDT)), Multi-Configurational Self-Consistent Field (MCSCF) methods (Complete Active Space Self-Consistent Field (CASSCF), Complete Active Space Perturbation Theory of second order (CASPT2)). The other kind is electron-density-based methods, such as time-dependent density functional theory (TDDFT) method. In recent years, CASSCF and TDDFT methods are widely used to deal with the photochemical properties from small to medium molecules and even large biological systems. In the next sections, we will focus on these two popular computational methods for excited states.

2.2.1 Time-Dependent Functional Theory (TDDFT)

To solve the Schrödinger’s equation, traditional quantum chemical approaches such as \textit{ab initio} methods, semi-empirical methods depend on the determination of the wavefunction.\(^{42}\) When the number of electrons increases, the wavefunctions become much more complicated and it will cost more computing time. However, in density functional theory, the electron density of a molecule is used to determine the energy and derivative properties of molecules.\(^{43-45}\) The electron density only depends on three spatial coordinates. It is a function with three variables: x-position, y-position, and z-position of the electrons (\(\rho(x, y, z)\)). The energy of the molecule is then a functional of the electron density.

\[
E = F[\rho(x, y, z)]
\]  

(2.1)

So there is a one-to-one mapping between the electron density of a system and the energy. We can get numerous information of a molecule if we can determine its electron density which is the basis for density functional theory. Compared to \textit{ab}
initio methods, the electron density function is only dependent on three coordinates, independently of the system size. This approach is much faster than ab initio methods.

To look back at density functional theory’s history, from the Thomas–Fermi model\(^{46}\) (developed by Thomas and Fermi in 1927) to Hohenberg-Kohn theorem\(^{47}\) which is thought to be the real birth of density functional theory, DFT has come a long way.

In 1964, Hohenberg and Kohn set up the famous Hohenberg-Kohn theorem.\(^{47}\)

The first Hohenberg-Kohn theorem ensures that the external potential is uniquely determined by the electron density of the ground state. The second Hohenberg-Kohn theorem guarantees the existence of a variational principle for electron densities.

Kohn and Sham introduced the Kohn-Sham orbitals and developed the Kohn-Sham theory which brings the density functional theory into a more practical version.\(^{48}\) In Kohn-Sham theorems, the total energy \(E[\rho]\) is shown as follows:

\[
E[\rho] = T_\text{s}[\rho] + J[\rho] + E_\text{ne}[\rho] + E_{\text{xc}}[\rho]
\] (2.2)

Where, \(T_\text{s}[\rho]\) is the kinetic energy of a non-interaction system, \(J[\rho]\) is the classical electron-electron repulsive (Coulombic) energy, \(E_\text{ne}[\rho]\) is the nuclear-electron attraction energy. The exchange-correlation term \(E_{\text{xc}}[\rho]\) contains the exchange and correlation effects. In density functional theory, all the approximations lie in the exchange-correlation term.

The Kohn-Sham equation:

\[
\hat{H}_{\text{KS}} \phi_i^{\text{KS}} = \varepsilon_i \phi_i^{\text{KS}}
\] (2.3)

Where, the Hamilton \(\hat{H}_{\text{KS}}\) can be expressed as:

\[
\hat{H}_{\text{KS}} (r) = -\frac{1}{2} \nabla^2 + V_{\text{KS}} (r)
\] (2.4)

Where,

\[
V_{\text{KS}} (r) = V_{\text{ne}} (r) + \int \frac{\rho(r')}{|r-r'|} dr' + V_{\text{xc}} (r)
\] (2.5)

\(V_{\text{xc}} [r]\) is the exchange-correlation term. For present-day DFT theory, the
exchange-correlation functionals can be divided into three types: local spin density approximation (LSDA) functionals,\textsuperscript{49,50} generalized gradient approximation (GGA) functionals,\textsuperscript{51-53} and hybridized functionals.\textsuperscript{54,55}

However, equation (2.3) is limited to time-independent systems. The Runge-Gross Theorem\textsuperscript{56} is an analogous time-dependent version of the first Hohenberg-Kohn theorem. It states that two external potentials \( v(r_1,t) \) and \( v(r_2,t) \) differed by more than a time-dependent constant \( C(t) \) result in two different electron densities, that is:

\[
v(r_1,t) \neq v(r_2,t) + C(t) \Rightarrow \rho(r_1,t) \neq \rho(r_2,t)
\]

So there still exists a unique relationship between time-dependent potentials \( V(r,t) \) and time-dependent densities \( \rho(r,t) \). Therefore the property of system can be written as a functional of the time-dependent density. The Runge-Gross Theorem is rigorous basis of TDDFT. This is the first step for the extension to the time-dependent domain. The next step is the existence of a time-dependent variational principle that is analogous to the second Hohenberg-Kohn theorem.\textsuperscript{30,57-58}

If wave function \( \Psi(r,t) \) is the solution of the time-dependent Schrödinger equation,

\[
i \frac{\partial}{\partial t} \Psi(r,t) = \hat{H}(r,t)\Psi(r,t)
\]

It is the stationary point of the action integral \( A \):

\[
A = \int_{t_0}^{t_f} dt \left( \Psi(t) \left( i \frac{\partial}{\partial t} - \hat{H}(t) \right) \Psi(t) \right)
\]

According to the Runge-Gross theorem, the action integral can be written as a functional of the time-dependent density:

\[
A[\rho] = \int_{t_0}^{t_f} dt \left( \Psi[\rho](r,t) \left( i \frac{\partial}{\partial t} - \hat{H}(r,t) \right) \Psi[\rho](r,t) \right)
\]

To derive the time-dependent Kohn-Sham equation, a time-dependent noninteracting reference system exists according to van Leeuwen. The noninteracting density is equal to the exact density of the real interacting system. The time-dependent Kohn-Sham equation is written as:
\[
\frac{i}{\hbar} \frac{\partial \phi_i(r,t)}{\partial t} = \left( -\frac{1}{2} \nabla_i^2 + v(r,t) + \int d^3r' \frac{\rho(r',t)}{|r - r'|} + \frac{\delta A_{xc}[\rho]}{\delta \rho(r,t)} \right) \phi_i(r,t) \quad (2.9)
\]

In this case, the \( A_{xc} \) part contains all the exchange and correlation effects. There are different approximations for this functional. The widely-known one is the adiabatic local density approximation (ALDA).

By now, the DFT theory is extended to the time-dependent domain and developed to time-dependent density functional theory (TDDFT). The TDDFT in linear response is thought to be another milestone of TDDFT theory. Using the linear response of the time-dependent Kohn-Sham equation, we can obtain useful information, such as excitation energies and oscillator strengths of excited states.

TDDFT has become one of the most popular quantum chemical tools from first principles to calculate excited-state properties of medium-sized or even large biological molecules. Since accurate high level calculations such as CASPT2 and CASSCF employing large active spaces are tedious and time consuming; the TDDFT calculation can reach the accuracy of sophisticated quantum chemical methods with moderate computational cost. However, there are shortcomings in TDDFT. The chosen of the right exchange-correlation functional for the given excited state property is crucial. The hybrid functional B3LYP is unsuccessful in some applications. For instance, the lack of the long range correlation causes the charge transfer (CT) problem. We did TDDFT calculations for 2-aminopyridine dimer with different functionals. The results are compared in Figure 2.3. Coulomb-attenuated hybrid exchange correlation functional CAM-B3LYP\(^{59-60} \) is specially designed for treating the long-range CT transitions. Compared to the CAM-B3LYP functional, one can see that the TDDFT/B3LYP underestimates the energies of CT states which are shown in red color in Figure 2.3.
2.2 Popular Computational Methods for Excited States

2.2.2 Complete Active Space Self-Consistent Field (CASSCF) method

First of all, we give a brief introduction of the Configuration Interaction (CI) method and Multi Configurational Self-Consistent Field (MCSCF) method.\textsuperscript{40,42}

Configuration Interaction method uses molecular orbitals based on a Hartree-Fock calculation. The wave function is a linear combination of HF wave function and other electronic configurations. The contributions of excited states are included in a linear variation function for the ground state.

\[
\langle \Psi_{CI} \rangle = \{ \Psi_0 \}, \{ \Psi_f \}, \{ \Psi_{ab} \}, \{ \Psi_{abc} \}, \cdots
\]

\(| \Psi_f \rangle, | \Psi_{ab} \rangle, | \Psi_{abc} \rangle\) are single(S), double(D) and triple (T) configurations respectively.

The CI wavefunction can be represented as:

\[
| \varphi_{CI} \rangle = C_0 | \Psi_0 \rangle + \sum_{\iota, \alpha} C_{i\alpha} | \Psi_{i\alpha} \rangle + \sum_{i>j, \alpha\beta} C_{ij\beta} | \Psi_{ij\beta} \rangle + \sum_{i>j>k, \alpha\beta\gamma} C_{ijk\gamma} | \Psi_{ijk\gamma} \rangle + \cdots
\]

Full configuration interaction (Full-CI) includes all the possible excitations from the HF ground state to all the virtual orbitals. However, full CI is computationally extremely expensive because of the large majority of the basis sets and configurations. So it is only possible for small systems and small basis sets. To put that into a more practical term, one selects important configurations to use the truncated CI method.
For instance, Configuration Interaction with Single Excitation (CIS) is one of the easiest ways to get excited state energies. It starts from the Hartree-Fock wave function and promotes one electron to one of the virtual orbitals. There are also other methodologies considering the multi-configurations interactions, such as CISD (Configuration Interaction with Single and Double excitations) and CISDT which includes the triple excitations.

The CI calculations take the Hartree-Fock ground state wavefunctions as the reference states of molecules. If it encounters the case that the Hartree-Fock determinants are not adequate, e.g. when the coefficient $C_i$ is very small, we need to choose more electronic configurations for the reference states. Multi-configurational self consistent field (MCSCF) calculations are used as an improvement of configuration interaction method. The wavefunction of MCSCF method $\Psi_{\text{MCSCF}}$ is a linear combination of different configuration state functions (CSFs) which is written as $\Phi_k$ in equation 3.2.

$$\Psi_{\text{MCSCF}} = \sum_k A_k \Phi_k$$  \hspace{1cm} (3.2)

Where $\Phi_k$ is again written as determinants of molecular orbital ($\varphi_i$) which is linear combination of the atomic orbital ($\chi_\mu$) also.

$$\Phi_k = \frac{1}{\sqrt{N!}} \text{Det} \left[ \prod_{i=1}^{N} \varphi_i \right]$$  \hspace{1cm} (3.3)

$$\varphi_i = \sum_\mu \chi_\mu C_{\mu i}$$  \hspace{1cm} (3.4)

So in the MCSCF calculations, the coefficients of both the configuration state functions and the basis functions in the molecular orbitals are optimized so as to minimize the energy. The MCSCF method has faster convergence than that of CI and MCSCF wavefunctions are widely used as reference states in Multireference configuration interaction (MRCI), complete active space SCF (CASSCF) method, or complete active space with second-order perturbation theory (CASPT2) to calculate the excited states’ properties.
The Complete Active Space Self-Consistent Field (CASSCF) method is a particularly widely used approach of MCSCF and it was first propounded by Ruedenberg and Roos.\(^6\) It is one of the few popular methods that are suitable for the optimizations of both the ground states and excited states. Similar to other MCSCF methods, the most important problem in CASSCF is how to choose the configurations. The orbitals used in CASSCF situation are usually classified into inactive orbitals and active orbitals. The inactive orbitals include the so-called core orbitals which are doubly occupied and also the virtual orbitals. The active orbitals are the most important orbitals involved in the reactions of the systems. Electrons that occupy active orbitals called active electrons. Usually the number of the electron occupation in active orbital is in the range from 0 to 2. Then the active electrons and active orbitals constitute the active space. The CASSCF wavefunction consists of a linear combination of all the configurations that arise from distribution of the active electrons among active orbitals under given spatial and spin symmetry. For example, the CASSCF wave function of active space CAS (10, 8) is built by distributing of 10 active electrons in 8 active orbitals. That is to say, 10 active electrons are distributed between all configurations that can be constructed from 8 active orbitals. The active space is the key concept of the CASSCF method. Different orbitals and electrons in the active space may have significant impact on the relative and absolute energy of the studied system. Generally, the active space should include all the valence orbitals and valence electrons. However, along with the expansion of the active space, the computational cost increases quickly. The larger active space is, the more computational resource costs. So in practical application, we only select relevant orbitals and electrons according to the chemical reaction process that we are interested. In contrast to other MCSCF approaches, the CASSCF method avoids the randomicity of the composition of the configuration functions and involves no selection of individual configurations. It is reasonably accurate for the equilibrium structures and properties of the ground states and electronic excited states under appropriate basis sets. Meanwhile, the CASSCF method can also give good description for the transition states and crossing points between different potential energy surfaces.
Therefore, it is frequently used to solve the complex photo-induced chemical reactions that always involve the molecular excited states.

However, the CASSCF method also has some limitations. The CASSCF method itself only includes a small fraction of the electronic correlation energy and does not include electronic dynamic correlations outside the active space. So there is error in the electronic energies. Therefore, if we want to get more accurate excitation energy, we need to include correction to dynamic correlation. Starting from the CASSCF wave functions, there are some ways to consider the dynamic correlations, such as MR-CI, MRPT (MR-MP2 and CAS-PT2) and MRCC approaches. An overall protocol named CASPT2//CASSCF is widely used. What’s more, the CASSCF method is limited by the system’s size. Especially in biological systems, due to the large number of atoms, the CASSCF method is always combined with classical molecular mechanics (like QM/MM method) to deal with the chemical reactions in proteins.
Chapter 3

Deoxyribonucleic acid (DNA) and its photochemistry

3.1 The Building Blocks of DNA

![Diagram of nucleic acid and deoxynucleotide structure]

**Figure 3.1** The basic components of a nucleic acid and a deoxynucleotide

Nucleic acids are the basic biological macromolecules of heredity. Their structural unit is nucleotide, consisting of phosphoric acid, pentose sugar (ribose or ribodesose), and base (purine or pyrimidine) as shown in Figure 3.1. According to the different types of the pentose sugar, nucleic acids are classified into the ribonucleic acids (RNAs) and deoxyribonucleic acids (DNAs). In RNA, the pentose sugar contains a hydroxyl group at the 2’ carbon atom of the five membered-ring. In DNA, the 2’-OH is replaced by hydrogen as shown in Figure 3.1. The genetic code in DNA is mainly stored in four bases: adenine (A), thymine (T), cytosine (C), and guanine (G), see Figure 3.2.
3.2 DNA double helix

Each base of DNA is linked to the backbone, composed of sugars (2-deoxyribose in DNA) and phosphates. And repeated units of deoxynucleotide constitute the polynucleotide, which is termed as one DNA strand. Because the sugar units is bonded with phosphate groups by the formation of ester bonds between C3’ and C5’ of adjacent sugar rings, the terminals of DNA strand are often defined as 3’ and 5’ terminals.

In 1953, Watson and Crick discovered the famous double helix structure of the DNA (see Figure 3.3). In this structure, the double-strand DNA is built up by two anti-parallel, right handed strands running in opposite directions along the same axis. The two adjacent base pairs are in average separated by 3.4 Å and a complete turn covers 10 base pairs with ca. 34 Å, while the radius of the helix is 10 Å.

The four main DNA bases pair up with each other following the “base-pair rule” which states that A only bonds with T and C only bonds with G. See the base pair A-T and G-C in Figure 3.3. In base pairs, there are hydrogen bonds between different
3.2 DNA double helix

bases (A-T: two hydrogen bonds, G-C: three hydrogen bonds) which connect the two DNA single strands together.

![Diagram of DNA double helix and base pairs A-T and C-G.]

**Figure 3.3** The structures of DNA double helix and base pairs A-T and C-G.

The discovery of the double helix structure of DNA is thought to be one of the most significant milestones in the human science history. After that, the DNA science ranks as one of the most popular fields of the last few decades, and it will continue to be the frontier science since it leads the way to a better understanding of the basic units of human being.
3.3 DNA photostability

As the genetic material of living organisms, DNA is somehow vulnerable and can easily be damaged by external perturbations such as UV radiation from the sun.\textsuperscript{13} The DNA bases are good chromophores that have intense absorptions at UV-light around 260 nm.\textsuperscript{62} The DNA nucleobases can be promoted to the excited states by absorbing radiation energy. The excess energy needs to dissipate via several pathways to the ground states, or it will initiate photochemical reactions that finally results in the formation of harmful photolesions and bring out mutations and carcinogenesis. Fortunately, the quantum yield of photolesion formation is less than 1%.\textsuperscript{15} This indicates that most excited DNA molecules just relax back to their ground states, avoiding harmful photochemical reactions which lead to photo-damaged lesions. In this sense, DNA has its own resistance to UV-light induced photochemical damage. This property is called photostability of DNA, which is the potential choice for biological evolution and is major topic of debates for decades.\textsuperscript{15}

As mentioned above, starting from the single bases which are the building blocks of DNA, base pairing and base stacking and other interactions ultimately constitute the double helix structure. There is therefore great interest in which effect leads to the photostability of DNA.

![Figure 3.4 Several possible factors of DNA photostability](image-url)
Figure 3.4 represents some suggested effects for the photostability, which will be discussed in following paragraphs. For the DNA base monomers, relative low fluorescence quantum yields (about $10^{-4}$) in aqueous solution demonstrate that there exists a highly efficient nonradiative decay pathway after UV-light absorption.\textsuperscript{63-64} Experimentally, in 2000, the first accurate ultrafast femtosecond transient absorption measurements showed that the $1\pi\pi^*$ states of DNA and RNA base monomers decay to the ground state within a subpicosecond or even femtosecond timescale.\textsuperscript{65} At present, there is good consensus between experimentalists and theoreticians that the ultrafast nonradiative decay of DNA single base is mostly caused by efficient internal conversions through conical intersections (CI).\textsuperscript{65-68} Theoretically, conical intersections have been found in almost all natural DNA nucleobases (A, T, C, G) and many other of their derivatives.\textsuperscript{67} The conical intersections in different bases are of different characters, such as $1\pi\pi^*/S_0$, $1\pi\sigma^*/S_0$, $1\pi\pi^*/1\pi\pi^*$, $1\pi\pi^*/1\pi\sigma^*$, and so on. These low-lying CIs or even high-lying CIs can be accessed from the Franck-Condon region through nearly barrierless paths, and favorable internal conversion from the excited singlet states to the ground state happens. The diverse CIs are usually possessed of typical ring deformation structures that involve in out-of-plane motions.\textsuperscript{67,69}

**Figure 3.5** Structure of 2-aminopyridine (a) and the energy profile for the ring deformation reaction (b). Selected from Paper II, reprinted with permission from American Institute of Physics.
We use *ab initio* methods and RRKM theory to study the radiative and nonradiative decay of 2-aminopyridine which is a frequently used model system for nucleobase, see Figure 3.5 (a).

Compared to other possible nonradiative photochemical pathways (like aromatic ring opening, intramolecular hydrogen transfer, hydrogen detachment), our calculations shown that an ultrashort nonradiative ring deformation pathway on *S*\(_1\) state is the most efficient nonradiative deactivation for 2-aminopyridine.

We have located the conical intersection between the *S*\(_1\) excited state and the ground state, denoted as *S*\(_1\)/*S*\(_0\) in Figure 3.5(b). One carbon atom in the aromatic ring is out of plane, accompanied by twisting of the C-C and C-N bonds. From the energy profile in Figure 3.5(b), we can see that when the molecule is excited upon to the first excited state *S*\(_1\), it can access the conical intersection *S*\(_1\)/*S*\(_0\) via a low energy barrier along the ring deformation coordinate. At the same time, the *S*\(_0\) state rises sharply and finally it meets the surface of *S*\(_1\) state at the point *S*\(_1\)/*S*\(_0\). It is explained by Kohler *et al.* that in the ring deformation structure, the aromaticity is destroyed.\(^{15}\) The ground state energy increases sharply due to the loss of \(\pi\)-bond stabilization; while the excited-state energy is relatively insensitive to the ring deformation.\(^{15}\)

In summary, the conical intersection plays an important role in the nonradiative decay pathways for the single base. By means of this efficient internal conversion, the excited monomers can release their initial energy and relax to the ground state avoiding injurious photodamage. Some excellent reviews emphasized again that the CIs are responsible for the ultrafast internal conversion in different kinds of single base monomers.\(^{14,67,70}\) Bern Kohler and co-workers stressed the nonradiative decay mechanisms in DNA from the experimental point of view.\(^{14,70}\) In the meantime, theoretical calculations on DNA natural bases and tautomers emphasized the importance of the CIs by using accurate quantum chemical approaches.\(^{67}\) Those studies indicate that the special photochemical feature of DNA nucleobases may attribute to the natural selection of the molecular evolution.

Different bases can assemble together horizontally (base pairing) or vertically (base stacking). These spatial organizations may create further new deactivation passages.
For instance, base pairing effect originates from the hydrogen bonds between single bases. Domcke et al.\textsuperscript{71} used Femtosecond time-resolved mass spectroscopy and detected an excited state of $65 \pm 10$ ps for the 2-aminopyridine dimer which is a model system for the nucleic acid base pairs, as shown in Figure 3.6 (a).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.6.png}
\caption{The electron-driven proton-transfer process in (a) 2-aminopyridine dimer and (b) Adenine-Thymine base pair. Selected from paper III, reprinted with permission from American Institute of Physics.}
\end{figure}

Their later \textit{ab initio} calculations proposed that the conical intersection between low-lying $^1\pi\pi^*$ local excited state (LE state in Figure 3.6(a)) and a charge transfer state (CT1 and CT2 states in Figure 3.6(a)) along the N-H proton transfer coordinate was responsible for the observed short lifetime. This electron-driven proton-transfer process was also revealed in the A-T and G-C base pairs.\textsuperscript{72,73} As seen in Figure 3.6 (b), after excitation to the LE state of the A-T base pair, the molecule can overcome small barrier and reach the crossing seams between the LE state and CT states. There are different kinds of charge transfer states and the CT state of A$\rightarrow$T$^*$ type is energetically
more favorable than that from T to A, denoted as CT (T→A) in Figure 3.6 (b). The charge transfer from adenine to thymine leads to spontaneous proton transfer in the same direction to neutralize the charge separation. Along the proton transfer coordinate, the energy of ground state increases while the CT states are essentially repulsive. At longer N-H distance, internal conversion will occur and finally the excited molecule decays back to the ground state. As has been shown above, the electron-driven proton charge-transfer states always play a decisive role in the relaxation processes of the base pairing system. There are also experimental studies which support this light-induced proton transfer mechanism.\textsuperscript{71,74,75} However, questions are raised when the base stacking effect is included into the oligo- and polynucleotides that makes the situation more complicated.

The base stacking effect is a significant interaction to the stability of DNA. It includes overlapping of π orbitals, dispersion attraction and electrostatic interactions, and so on. Since the nucleobases possess the purine or pyrimidine rings which are suitable for stacking, the base stacking generally exists in the single- and double-stranded DNA. For DNA oligo- and polynucleotides, in addition to the fast nonradiative decay rate in those single bases, experimental evidence of one or two orders of magnitude slower rate also exists.\textsuperscript{76} It suggests that there are long-lived states which compete with the ultrafast monomer-like decay channels. Crespo-Hernández and Kohler studied the single-stranded (dA)\textsubscript{18} \textsuperscript{76}, and indicated that there were both ultrafast (τ≈ 1 ps) and a slower (τ≈ 126 ps) components exist in the single strand. It’s surprising that same kinetics has been observed in the duplex (dA)\textsubscript{18}·(dT)\textsubscript{18}. So they concluded the characteristic long-lived signals in DNA oligo- and polynucleotides were ascribed to the stacking effect instead of the base pairing. This work thereby evokes a series of controversy about the primary and secondary roles of the base pairing and base stacking effects.\textsuperscript{14,15,70,77} Besides, there are other possible explanations for the slow nonradiative decay in DNA oligo- and polynucleotides.

Some researchers concluded that the excimer, exciplex or excitonic states formed by a pair of stacked bases are responsible for the long-lived signals.\textsuperscript{15} Others thought
that it was because of the delocalized excited states, such as Frenkel excitons or interbase charge-transfer states, exist in the DNA single and double strands.\textsuperscript{70, 78, 79} But the delocalization length is unknown. In Figure 3.7, we illustrate the detachment (blue color) and attachment (red color) electron densities calculated at PCM/TD-B3LYP for the excited singlet state $S_{11}$ for the A$_3$ and T$_3$. It is obvious that the $S_{11}$ state of A$_3$ displays much delocalized property while the $S_{11}$ state of T$_3$ is a charge transfer state. What’s more, interstrand proton transfer and other channels via dark states like $^1\text{nn}\pi^*$ states have also been suggested.\textsuperscript{70} However, there are lots of arguments and great uncertainty about the nonradiative decay in the above mechanisms. This is not only because of the complexity of the DNA double helix itself but also the intricate biological environment around the real DNA molecules. Since the underlying nonradiative decay mechanisms of the DNA oligo- and polynucleotides are poorly understood, more challenges and opportunities in this active topic will continually make it a subject of great interest both experimentally and theoretically.

\textbf{Figure 3.7} Detachment (blue) and attachment (red) electron densities for the selected $S_{11}$ ($\pi\rightarrow\pi^*$) states for A$_3$ and T$_3$. 

- 24 -
3.4 DNA photodamage

The photostability of DNA can prevent most of the photodamage formations in DNA. However, once a chemical bond in one of three billion nucleotides of human being is modified, the outcome will be catastrophic. From this perspective, studies on DNA photodamage have great practical significance in the DNA photochemistry. Since the skin cancer induced by UV radiation has drawn much attention these years because of its high incidence rate, the photodamage appears to be crucial in photocarcinogenesis. UV exposure results in a variety of potentially mutagenic cellular lesions, referred as photoproducts. Exposed by carcinogenic far-UV wavelengths brings in three characteristic lesions, resulting in cyclobutane pyrimidine dimers (CPD), pyrimidine (6–4) pyrimidone photoproducts [(6–4)PP] and their Dewar valence photoisomers (DewarPP).\(^{80-82}\)

The CPDs come from the covalent linkage between two adjacent pyrimidines by the formation of a four-membered ring structure, see Figure 3.8, the C5-C6 double bonds of pyrimidines are linked and saturated in CPDs. CPDs are the most abundant photoproducts of all three photoproducts. It is said that the CPDs and the (6–4)PPs constitute around 75% and 25% of the DNA photodamage products, respectively.\(^ {18}\) Since described as early as 1960, the reaction and repair mechanism of CPDs have been extensively studied by means of both experimental techniques and theoretical computations.\(^ {23,83-85}\) Now, it is well recognized that the mutagenic and carcinogenic potential of UV-induced damage is expected to arise mainly from CPDs. Studies have shown that CPDs are the main trigger for apoptosis in nucleotide excision repair (NER)-proficient cells.\(^ {23}\)
Figure 3.8  Structures of three DNA photoproducts. Selected from paper IV, reprinted with permission from American Chemical Society.

Different from the CPDs, as illustrated in Figure 3.8, an incorrect covalent linkage between the C6 atom of the one pyrimidine and the C4' atom of the adjacent pyrimidine results in the (6-4)PPs. Although smaller proportion than CPDs, in some cases, the (6-4)PPs are predicted to be more potentially lethal than CPDs. For instance, in the NER-deficient cells, experimental data indicated that the (6-4)PP lesions were more toxic and mutagenic compared with CPDs, which made it currently a subject of intensive studies.

When exposed to UV-A/B light, the formation of the N3'-C6' bond in the pyrimidone ring leads to its valence isomer called DewarPP. The quantum yield of DewarPP is quite low and this hampers the study of experiments and theoretical calculations on this naturally formed photoproduct. So compared to the other two photoproducts, the DewarPP has been less studied. The cytotoxic and mutagenic properties of this photoproduct are not well established to date. In particular, the
underlying mechanisms of the photoisomerization between the (6-4)PP and DewarPP remain unclear at the molecular level.

To sum up, the above three photoproduc.ts are considered to be responsible for mutation induction in critical genes, tumorigenesis, cell death, and so on. These will drastically impair the transcription and replication of DNA and finally lead to carcinogenesis like skin cancer. Therefore, if the photo-damaged DNA lesions cannot be repaired or removed, they will infect the normal cellular processing of DNA and cause mutations or even death.
3.5 DNA photorepair

In order to maintain the regular operation for the biophysical activities, especially for the genetic stability and continuance of life, the organisms have a set of restoration system to repair specific damaged lesions. Especially for DNA photodamage, there are several approaches for organisms to repair the UV-induced photo-lesions. First, the excision repair, which includes the base excision repair (BER) and nucleotide excision repair (NER), is one kind of common repair mechanisms.88 In this mechanism, the damaged DNA is replaced with new nucleotides by cutting off and then interlinking nucleotides in the damaged part. Second, using the energy of light, many organisms can undergo photoreactivation process which involves in particular enzymes called CPD photolyase and (6-4)PP photolyase.13,88 The photoreactivation refers to the reversal process of the photodamage reactions by the enzyme upon irradiation with blue or near-UV light (wavelength 300-450 nm).13 The DNA photolyases are monomeric enzymes found in a variety of organisms including higher plants and bacteria. The photolyases have 450-550 amino acids in length, with the total molecular weight of 55-65 kDal.88 The X-ray structure of the CPD photolyase found in Escherichia coli (PDB entry:1TEZ) is shown in Figure 3.9. The photolyase has two domains which are labeled as AB and A domains in the N-terminal and C-terminal, respectively. In A domain, the enzyme folds up a spatial cavity which is the binding position for the damaged DNA lesions and cofactor. In Figure 3.9, the photolyase binds to the oligomer DNA (in green) duplex with a photodamaged CPD flipped out and inserted into the active site.

There are two kinds of important chromophore cofactors in the photolyase. First is so-called the light-harvesting cofactor, which can be either a methenyltetrahydrofolate (MTHF) or an 8-hydroxy-5-deazariboflavin (HDF), as colored red in Figure 3.9. People gives them a lively description — light antenna that absorbs a photon and transfers the energy to the other redox active cofactor and the repair rate will increase by ten to hundred-fold under light limitation by these light-harvesting antenna.13
Figure 3.9 Structure of DNA photolyase bound to DNA with a CPD. (PDB code:1TEZ) The solvated water molecules are not shown for clarity.

The second cofactor is the flavin adenine dinucleotide (FAD) which is shown in blue in Figure 3.9. The DNA photolyases are thus also called flavoproteins. There are three different oxidation states in the flavin molecules: the oxidized form (flavoquinone, FAD), radicals (flavosemiquinone, FADH’ and FAD’-), and the two-electron-reduced flavin-adenine dinucleotide (flavohydroquinone, FADH₂ and FADH’). We present the oxidation-reduction cycle in Figure 3.10. The redox reaction is also accompanied by the electron and proton transfer processes between different redox states. The flavins thus have indispensable functions in the enzyme repair mechanism of the DNA photodamage.
It is now well recognized that in the photorepair process, the catalytic active form in vivo is the anionic hydroquinone (FADH\(^{-}\)), see Figure 3.11. We will discuss its significant role in the photorepair mechanism of photolyase in the following chapter.

**Figure 3.10** The three redox states in the oxidation-reduction cycle of flavins.

**Figure 3.11** Structure and labeling of the FADH\(^{-}\) molecule. Selected from paper V, reprinted with permission from American Chemical Society.

We summarize the photoreactivation process for CPD lesions in Figure 3.12. In the first step, the CPD photolyase “recognizes” the damaged CPD dimer and binds to it
by twisting the CPD dimer out of the DNA normal oligomer into the active pocket.\textsuperscript{13} Then, the cofactor MTHF (or HDF) absorbs photon from the UV light and transfers energy to excite the FADH\textsuperscript{+} molecule (FADH\textsuperscript{*}) through a Förster resonance energy transfer (FRET) mechanism.\textsuperscript{89} The excited FADH\textsuperscript{*} has a sufficiently low redox potential to transfer an electron to the thymine dimer with the formation of a radical anion that subsequently undergoes cycloreversion. This electron transfer process has been observed by Kao \textit{et al.} via femtosecond synchronization method and the electron transfer rate was measured to be in the range of $5.5 \times 10^9 - 3 \times 10^{10}$ s\textsuperscript{-1}.\textsuperscript{90} The photo-induced electron transfer from FADH\textsuperscript{*} to the thymine dimer triggers a [2+2] cycloreversion which leads to the monomerization of the thymine dimer. The splitting of thymine dimer forms a thymine and a thymine radical anion. The latter one donates an electron back to the semiquinone flavin radical (FADH\textsuperscript{·}) to regenerate FADH\textsuperscript{+} which is responsible for another catalytic cycle. The repair process was observed to be completed in 560 and 589 ps for \textit{E.coli} and \textit{A. nidulans} photolyase, respectively.\textsuperscript{90,91} The repair process involves light and thus the DNA photolyase is also called the "photo-driven" flavoprotein. The whole photoreactivation is a very complex process with both electron and energy transfer.

\textbf{Figure 3.12} Proposed reaction mechanism for the repair of thymine dimer by DNA photolyase.
The photoreactivation process for the (6-4)PPs in (6-4) photolyase is quite similar to the CPDs. However, after the photo-induced electron transfer from FADH', the successive electron-induced repair step for (6-4)PPs goes in different direction. In contrast to CPDs, numerous studies proposed that a thermal four-membered ring oxetane intermediate is formed in the repair mechanism of (6-4) PP. The subsequent ring opening of this intermediate then splits the (6-4)PP into two base monomers, as shown in the right column of Figure 3.13.

**Figure 3.13** Oxetane-mediated and the non-oxetane repair mechanisms for the (6-4)PPs.

However, some experimental and theoretical evidences supported a non-oxetane repair mechanism. Direct OH-transfer from 5’- to 3’-base avoiding the formation of oxetane intermediate was proposed by Domratcheva et al. It was suggested that electron transfer occurs directly from protein to lesion but not to a strained oxetane intermediate as previously postulated. In this mechanism, the hydroxyl group of the pyrimidone transfers directly as depicted in Figure 3.13. However, the transition state...
has relative high energy. A barrierless non-oxetane pathway via the conical intersection involved in the excited lesion radical anion was found theoretically by Yamamoto and co-workers.\textsuperscript{94} However, it was then questioned by Philipp that the (6-4) lesion radical anion was in the ground state instead of the excited state due to the insufficient absorbed photon energy.\textsuperscript{95} Besides, a transient-water-molecule-formation model was suggested by several groups and the two adjacent histidine residues were suggested to play important roles in the proton transfer process.\textsuperscript{96-98}

\textbf{Figure 3.14} Model for the study of (6-4)PP repair mechanism constructed from crystal structure (PDB code: 3CVU).

In order to reveal the underlying repair mechanism, we check the above suggested mechanisms based on the cluster model (Figure 3.14) by using the cluster approach\textsuperscript{99-105} at the B3LYP/6-31G(d) level, where the photoproduct T(6-4)T is colored mauve in the active site. Several functional groups around the photo lesion are included. Two histidines (green) and a water molecule WAT369 which is quite close to the pyrimidine ring, are chosen in the cluster model. The flavin FADH\textsuperscript{−} is simplified by the Adenine and 2’-OH. Other residues are collected to account for the strong hydrogen bonding network. It is important to mention that in order to check the
role of the residue His365 which is suggested to be the proton donor in the repair mechanism, we treat His365 in both neutral (N) and protonated (P) forms, labeled as HIE365 and HIP365, respectively. Larger basis set 6-311++G(2d,2p) is used in the single-point calculations based on the optimized structures. The polarizable continuum model (PCM) with the dielectric constant of 4 is used to consider the solvent effect and the rest part of the enzyme. 100-105

We first discuss the model with HIE365 inside. The optimized structures and energy profiles are illustrated in Figure 3.15, where the reactant with HIE365 is represented as N-react. The WAT369 forms a strong hydrogen bond (1.83 Å) to 5-hydroxyl group and it is further stabilized by the Y306 residue with a hydrogen bond (distance of 1.80 Å). Meanwhile, there is also a hydrogen bond between HIE365 and pyrimidone. The HIE365 and H369 are stabilized by 2’-OH and Y234 by hydrogen bonds, respectively. Starting from this reactant, we obtain the transition state for the non-oxetane repair mechanism which involves the direct hydroxyl transfer process. We investigate the stepwise hydroxyl transfer process (denoted as path A). The C5-O bond is increased to 1.85 Å in the transition state N-A-TS1 where the hydroxyl group is stabilized by HIE365 and WAT369. However, the barrier for the C-O cleavage is 31.6 kcal/mol at the PCM//B3LYP/6-31G(d) level. The hydroxyl group forms a hydrogen bond with His369 in following intermediate N-A-int which is 3.2 kcal/mol above the N-react. The next nucleophilic attack on the C4’ atom of pyrimidone produces a transition state N-A-TS2, with a reaction barrier of 20.3 kcal/mol. The product of the hydroxyl group transfer is represented as int2 in Figure 3.15. The C4’-O distance is 1.49 Å and the C6-C4’ bond is weakened from 1.53 Å of N-react to 1.67 Å of int2, which implies that the linked pyrimidines are gradually separated during the hydroxyl group transfer.
**Figure 3.15** Optimized geometries depicted by XYZViewer and relative energies for stationary points and transition states calculated at B3LYP/6-311++G(2d,2p) and PCM//B3LYP/6-31G(d) levels (in parenthesis) of HIE365 cluster model.
While for the cluster with HIP365 inside, the optimized structures and calculated potential energy profiles are summarized in Figure 3.16. We found the hydroxyl transfer process in a concerted way (denoted as path B in Figure 3.16) or the water-formation way (denoted as path C). The concerted hydroxyl transfer results in a transition state with strained four-member ring (P-B-TS1). The C5-O and C4’-O distance are 1.67 and 1.91 Å, respectively. However, the calculated reaction barrier for this converted process is 49.2 kcal/mol, which is impossible to be overcome in enzyme environment. On the other hand, the stepwise transfer of hydroxyl group induces the formation of a water molecule. After the cleavage of C5-O bond, the HIP365 gives the proton to the dissociative hydroxyl group which forms a water molecule in P-C-int1. This water molecule then attacks the C4’ atom of pyrimidone through transition state P-C-TS2. During this process, the attached proton returns to the histidine again. In this mechanism, the protonated histidine acts as a proton donor. Similarly, the calculated barrier 54.1 kcal/mol for the water-formation mechanism is also too high, as shown in Figure 3.16. Therefore, we exclude these repair pathways. Our calculations show that both the non-oxetane and the water-molecule-formation mechanisms have relatively high reaction barriers and the repair for the (6-4)PP may mainly undergo the oxetane mechanism.
Figure 3.16 Optimized geometries depicted by XYZViewer and relative energies for stationary points and transition states calculated at B3LYP/6-311++G(2d,2p) and PCM/B3LYP/6-31G(d) levels (in parenthesis) of HIP365 cluster model.
It is deserved to mention two recently published papers concerned from the standpoint of experiment and theoretical calculation, respectively. The former is by Li et al.\textsuperscript{106} who performed the ultrafast spectroscopy and found that the proton transfer process from His364 to (6-4)PP\textsuperscript{−} in the anionic ground state is the key step of the repair mechanism. The latter QM/MM study by Keyarash and co-workers suggested a new repair mechanism in which two photons were required together with the charge recombination steps.\textsuperscript{107} The calculated barrier was in good agreement with that observed by Li \textit{et al.}\textsuperscript{106} The oxetane-intermediate was also involved in their new repair mechanism.

\textbf{Figure 3.17} Depiction of isomerization process between (6-4)PP and DewarPP in (6-4) photolyase.

Compared to the CPD and (6-4)PP, the repair mechanism of DewarPP is still poorly understood to date. The repair of DewarPP involves not only transfer of the amino group to the C4’ atom, but also a N3’-C6’ bond breaking process which results
Chapter 3 Deoxyribonucleic acid (DNA) and its photochemistry

in the isomerization of DewarPP to (6-4)PP. The two kinds of C-N cleavage processes are competitive. Very recently, Glas et al.\textsuperscript{108} reported that the repair of DewarPPs by (6-4) photolyase involved first a rearrangement of the Dewar lesions into the corresponding (6-4)PPs via an electron injection process, see Figure 3.17. Later in our theoretical modeling, we found that the reaction barrier of isomerization from T(dew)T to T(6-4)T was sharply decreased from 30 (neutral ground state) to 15 kcal/mol (radical anion). It provides an efficient repair pathway for DewarPP, which well explained the experiments.\textsuperscript{108} We will discuss this in following chapter.
Chapter 4

Summary of Papers

4.1 Design of theoretical models for biological systems, *paper II, III, IV, V, VII*

In recent years, quantum computational methods have developed dramatically and have been successfully applied to investigate complex biochemical reactions. Theoretical modeling can describe underlying reaction mechanisms at atomic level and explain experimental observations. The primary task in theoretical modeling is to choose appropriate models for real biological systems. The selection of computational model for biomolecules needs to consider several aspects: the critical structural features, the surrounding environment like solvent effect, electrostatic effect, steric effect, or functional groups that participate in the reaction. In this section we will show some results concerning design of models for photochemical reactions that we are interested in.

As a simple model for the DNA nucleobase, 2-amiropyrindine (2-AP) which contains a N-H donor and an aromatic N acceptor group of the single base, is chosen to study the radiative and nonradiative decay pathways of DNA photostability in paper II. The doubly hydrogen-bonded dimer (see the structure in Figure 4.1) has been widely used as substitute for the unstable single base pairs in experiments, and a model molecule to mimic DNA base pairs in theoretical studies. 71, 75, 109
Figure 4.1 Structure of 2-AP dimer and DNA A-T base pair. Selected from paper III, reprinted with permission from American Institute of Physics.

However, compared to real DNA base pair like Adenine-Thymine (A-T) base pair shown in Figure 4.1, there are some basic differences between the model system and the real biological system. First, the 2-AP dimer molecule has a symmetric structure while A-T base pair not. Second, as mentioned before, the base pairs are covalently bonded to the sugar-phosphate backbone in the DNA double helix. However, this backbone constraint is missing in 2-AP dimer. So a question to be raised is that whether the 2-AP dimer can be a reasonable model for the A-T base pair?

In paper III, we use the time-dependent density functional theory to explore the similarities and differences in the ultrafast deactivation processes between 2-AP dimer and A-T base pair. The calculated two dimensional potential energy surfaces for the ground state of 2-AP dimer, A-T base pair with free and fixed N…N distances are shown in Figure 4.2. The two coordinates are the two main hydrogen bond distances as represented in Figure 4.1. One may conclude that the hydrogen transfer mechanisms for 2-AP dimer and A-T base pair are quite different in the ground states. The former exhibits the concerted hydrogen transfer mechanism while the latter shows the two-steps mechanism for the hydrogen transfer. So from this opinion, the 2-AP dimer is not a proper model for studying the ground state dynamics of the real DNA base pairs.
4.1 Design of theoretical models for biological systems

Figure 4.2 Two-dimensional potential energy surfaces of the ground state for (a) 2-AP dimer; and for the A-T base pair (b) with free and (c) fixed N…N distance. The picture is obtained from MATLAB 7.1. Selected from paper III, reprinted with permission from American Institute of Physics.

Based on the potential energy surfaces of the ground states, we get the corresponding excited state potential energy surfaces by TDDFT calculations as shown in Figure 4.3. Similar ultrafast deactivation pathways (pathway ① and ② in Figure 4.3) through the conical intersections between the LE (\( ^1\pi\pi^* \)) state and different electron-driven CT states have been found in both 2-AP dimer and A-T base pair. Although the height of barriers and the potential energy surfaces of excited states are different in these two systems, the calculated rate constants of the IC processes display the same trend.

Figure 4.3 Potential energy surfaces of the ground (S\(_0\)), charge transfer (CT), and localized (LE) states of (a) the 2-AP dimer; and the A-T base pair with (b) free and (c) fixed N…N distance. Two coordinates in the plane correspond to the main two hydrogen bridges, where the vertical coordinate is the energy in eV. Selected from paper III, reprinted with permission from American Institute of Physics.

Besides, from comparison of the potential energy surfaces for the A-T base pair with free and fixed N…N distance in Figures 4.2 and 4.3, it is clear that the reaction barrier is very sensitive to the intermolecular distance. The A-T base pair with frozen
N…N distance display much more steep potential energy surfaces and lower reaction barrier and in consequence the external force from the DNA backbone is one that actually exists and should not be ignored in the modeling.

Now it’s time to answer our question from the beginning, is the 2-AP dimer a reasonable model for the real DNA base pair? The answer “yes” depends on the particular purpose like kinetics of the IC process in the ultrafast deactivation pathway. Otherwise, the limited structural information will give incorrect results.

In paper V, we compare the widely used flavin model —7,8-dimethyl-10-methyl-isoalloxazine to the newly proposed model with intramolecular hydrogen bond by quantum chemical calculations and vibrationally-resolved electronic spectra simulations.

![Figure 4.4](image)

**Figure 4.4** Schematic diagram of the two flavin models. Selected from paper V, reprinted with permission from American Chemical Society.

The isoalloxazine ring (see M-NHD model in Figure 4.4) is the active core of flavin molecules and frequently used as a theoretical model in previous calculations. However, we found the existence of an intramolecular hydrogen bond between the 3’-OH of the ribityl moiety and the N1 atom of the isoalloxazine ring, see M-IHD model in Figures 4.4. The inclusion of this intramolecular hydrogen bond has an important impact on the geometrical and electronic structures of these two models, as
shown in Figure 4.5 and 4.6. In protein, the motion of the isoalloxazine ring is restricted by the strong intramolecular hydrogen bond in M-IHD and results in two stationary points in the ground state. Their “butterfly” bending angles are 14° and 9° respectively. However, with M-NHD model, we locate two degenerate states with a butterfly bending angle of 26° which is far from the experimental data (9°). In this sense, our new model is much more consistent with the experimental results. If we look at the structural information in solution, the intramolecular hydrogen bond in M-IHD is broken by a water molecule and thus the restriction is released. It seems that the intramolecular hydrogen bond of the flavin controls the flexibility of the isoalloxazine core in different biological environments. To have better understanding of these two models, we then compute the emission spectra for both models with a linear coupling model (LCM) method. The simulated results are shown in Figure 4.7.

**Figure 4.5** Geometrical and electronic structures of M-NHD model in protein (a,b,c) and in solution (d,e) obtained at CASSCF (10e, 8o)/6-31G(d) level. Selected from paper V, reprinted with permission from American Chemical Society.
Figure 4.6. Geometrical and electronic structures of M-IHD model in protein (a,b,c) and in solution (d,e) obtained at CASSCF (10e, 8o)/6-31G(d) level. Selected from paper V, reprinted with permission from American Chemical Society.

Figure 4.7 Calculated emission spectra of M-NHD and M-IHD models in the protein and in the solution environment. Selected from paper V, reprinted with permission from American Chemical Society.
If we compare our theoretical results with recent UV femtosecond laser spectrum by Zhong et al.\textsuperscript{110}, in Figure 4.7, the calculated dominate peaks of the new model in protein are at 510nm (515nm) and 545nm (545nm) which are in good agreement with experimental data in parentheses. Experimentally, upon 360nm excitation, the emission peak in solution is blue-shifted by 0.45eV compared to that in photolyase. This value is 0.30eV in our new M-IHD model. On the contrary, the conventional chosen model M-NHD fails in both the shapes and the energies of the spectra compared to the experiments. So from the spectroscopic point of view, it is clear to see that the experimental emission spectra are well reproduced by the M-IHD model and the presence of the intra-hydrogen bonding changes the main spectral features. We therefore suggest that this important intramolecular hydrogen bond needs to be considered when modeling the flavin molecules.

As a continuation of the work in paper V, in paper VII, the optical properties of three redox states of flavin mononucleotide (FMN) were investigated by comparing the calculated absorption spectra in different environment to experimental data. Flavin mononucleotide is the key cofactor in flavodoxin which has broad involvement in a variety of biochemical and physiological functions.\textsuperscript{111} There are three redox forms of FMN molecule which are shown in Figure 4.8.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.8.png}
\caption{Structure and labeling of the FMN in three redox states.}
\end{figure}

The absorption and emission spectra of these three redox forms exhibited quite different properties which were suggested to reflect local protein conformation plasticity and water network flexibility.\textsuperscript{112} The optimized structures of the intra-hydrogen bond models are illustrated in Figure 4.9.
Figure 4.9 Optimized structures for the three redox states of FMN, calculated at B3LYP/6-311++G(2d,2p) level.

The distinct difference in the geometric structure is the planarity of the isoalloxazine ring. The neutral and radical species possess relative planar ring while the FMNH\(^+\) has a butterfly bending angle of 149°. In addition, due to the negatively charged N1 atom in FMNH\(^+\), the intra-hydrogen bond between N1 and 2’-OH is much stronger. Based on these optimized geometries, we performed the TDDFT calculations and calculated the absorption spectra in both vacuum and PCM model, as shown in Figure 4.10.

Figure 4.10 Calculated vibrationally-resolved absorption spectra for three redox states of FMN in gas-phase and PCM model. Corresponding experimental spectra are compared in dark line.
One can conclude from Figure 4.10 that the external implicit-solvent environment has little influence to the spectra of neutral and radical species; while it significantly affects the spectra of fully reduced FMNH\(_2\) and it mainly arises from the intrinsic structural change and charge re-distribution during the redox cycle. We also set models to consider the specific environment, such as external hydrogen bond and surrounding protein environment. The calculated spectra matched well with our intra-hydrogen bond model and indicated that the external environment had little influence on the absorption spectra.

From above discussions, in this case, the biomolecules have the absolute initiative when facing the environmental changes, and their self-regulation acts critical role in their optical properties.

In addition to the internal structural characteristics, the external surrounding environments of the interested systems are another factor that needs to be considered in the design of computational models. The most common circumstances in biological reactions are protein and solution cases. In the present thesis, we consider these effects in several ways.

**Figure 4.11** structures and labeling of the computational models for the active site of DNA (6-4) photolyase.

In enzyme, the critical catalytic reaction takes place in the active site which is a
small region with functional groups. In Figure 4.11, we show the chosen model for the active site of DNA photolyase in paper IV and paper VI. It contains the main photo-damaged (6-4)PP lesion and the surrounding correlative residues. Usually, it is necessary to include the residues which take part in the catalysis or have strong interactions with the reaction centre (like hydrogen bond). And most of these residues are truncated to reduce the size. For example, residue Q299 is included because of the hydrogen bond with N3 and O2 atoms of (6-4)PP. Meanwhile, catalytic triad His365-His369-Tyr423 is served as a proton transfer chain in recent experimental studies on (6-4) photolyase. 97 The last but not the least, cofactor flavin adenine nucleotide (FAD) is truncated by the Adenine part and 2’-OH group which forms the hydrogen bond with His365. The rest protein surroundings are then treated as a homogenous polarizable medium with a dielectric constant of 4 (ε=4). This model is used to investigate the isomerization reaction and repair mechanism in (6-4) photolyase by the cluster approach methodology which has been widely and successfully applied in various enzymes. 99-105 Our calculations shown the protein environment has an important impact on the reaction barrier.

The solvent effect in present thesis is treated by different approaches. First is the Polarizable Continuum Model (PCM). The PCM model treats the solvent environment with a homogenous dielectric continuum medium of the dielectric constant ε based on the molecular-shaped cavity.113-117 This model is now one of the most popular approach to deal with the solvent effect.115

The second approach is the discrete model of QM/MM method. In paper IV, we built a water solvent box to mimic aqueous environment of photochemical reactions. The water box is calculated in classical molecular mechanical method. We summarized the calculated relative energies for (6-4)PP and DewarPP in Table 4.1. In aqueous solution, the barrier of the transition state (S_n-TS) of the N3’-C6’ bond cleavage in the S_n state is significantly decreased from 1 eV to 0.6 eV. The solvent molecules stabilize the transition state and open a new decay channel involving the “dark” excited states. So the solvent effect has a significant influence on the reaction mechanisms.
### Table 4.1 Calculated relative energies for (6-4) PP and Dewar PP in the ground and excited-states at CASPT2//CASSCF(10,8)/6-31G(d) level in gas phase and with CASPT2//CASSCF/Amber (QM/MM) protocol in solution.

In a word, as the initial start for the computational studies of biological reactions, the design of a proper molecular model is the primary work with several aspects considered. As the saying goes “Well begun is half done”, more detailed studies of appropriate model for different biological issues are required in the future.
4.2 Computational photophysical and photochemical studies on UV-induced DNA issues, *paper I* - *VI*

4.2.1 DNA photostability

As introduced in Chapter 3, the photostability of DNA is a decades-old question with obscurions and controversies. Single base, base pairing and base stacking *et al.* are suggested as several possible pathways for the deactivation of UV excited DNA molecules.

In *paper II*, we present that the fast internal conversion through the ring-deformation CI is responsible for the ultrafast nonradiative decay of lowest excited singlet state of 2-aminopyridine, which is a model system of DNA monobases. What’s more, we analyze the electron-driven proton charge-transfer deactivation pathway of the Adenine-Thymine base pair in *paper III*. In order to explore the effects of base pairing and base stacking in DNA photostability, in *paper I*, we employed model systems (dA)$_5$, (dT)$_5$, (dA)$_5$·(dT)$_5$ (see Figure 4.12) in combination with the density functional theory and time dependent DFT methods to investigate the excited state nature of DNA oligonucleotides.

![Figure 4.12](image)

**Figure 4.12** Computational models for DNA single strands and double strands. Selected from paper I, reprinted with permission from Wiley Periodicals, Inc..

In this work, the equilibrium structures of the ground-state are optimized at B3LYP (PBE0)/6-31G(d,p)/PCM level and the vertical excitation energies are obtained at TD-B3LYP (PBE0) and TD-LRC-BOP$^{118}$ (for CT states) level of theory. We present a
systematic study of the base stacking and hydrogen-bonding effects from single base (A, T) to base pair (A-T), single strand ((dA)_5, (dT)_5) and finally the double helix ((dA)_5-(dT)_5). The results are shown in Figure 4.13. On one hand, the absorption peak of first strong $\pi \rightarrow \pi^*$ transition for A and A$_3$ are 5.09 eV and 5.15 eV, respectively. The $^1\pi\pi^*$ excited states energies are increased due to the stacking interaction. However, energies of some other $^1\pi\pi^*$ excited states are reduced to 4.96 or 5.05 eV as shown in Figure 4.13. This explains the experimentally observed weak blue shift and red-side shoulder in the first strong absorption band. On the other hand, the H-bond interactions lead to a little decrease in energy from A (5.09 eV) to A·T (5.06 eV) and from A$_3$ (5.15eV) to (A·T)$_3$ (5.04eV). The same trend is also found from T$_3$ to (A·T)$_3$. The absorption intensity significantly decreased by the base pairing and stacking interactions from 0.13 in T$_3$ to 0.04 in (A·T)$_3$. In past decades, there are lots of controversies about the delocalization or charge-transfer nature of the excited state of DNA oligomers. By our work, we found that the redistribution of the $\pi$ electrons on excitation gives rise to a few delocalized partial charge-transfer (CT) excited states in A$_3$ but substantial CT characteristic excited states in T$_3$. These T→T charge transfer states which have low oscillator strengths produce the dark states in (A·T)$_3$. The A→T charge-transfer states are predicted to have relatively high energies (6.0eV-7.0eV) while the T→A transitions are of even higher energies.
Figure 4.13 The stacking and H-bonding interaction effects on $^{1}\pi\pi^*$ excited-state nature in DNA oligonucleotides. Blue lines: Electronic excitation localized on A bases; Green lines: Electronic excitation localized on T bases; Magenta lines: $T \rightarrow T$ charge-transfer excitations; Red lines: $A \rightarrow T$ charge-transfer excitations; Black line: Mixed charge-transfer excitation. Selected from paper I, reprinted with permission from Wiley Periodicals, Inc..

At last, we compared the base pairing and stacking effects on $^{1}\pi\pi^*$ states. As illustrated in Figure 4.14, the base pairing effect (increase in energy) is opposite to the base stacking effect (decrease in energy) on $^{1}\pi\pi^*$ states. The final result is that the relative energy of the lowest $^{1}\pi\pi^*$ state is almost the same from mono-nucleobases to double helix.

Figure 4.14 The stacking and H-bonding interaction effects on $^{1}\pi\pi^*$ excited states in DNA oligonucleotides. Blue lines: Electronic excitation localized on A bases; Magenta lines: Electronic excitation localized on T bases. Selected from paper I, reprinted with permission from Wiley Periodicals, Inc..
4.2.2 DNA photodamage

In paper IV, we studied the photo-isomerization reaction between two photo-damaged lesions: T(6-4)T (pyrimidine (6-4) pyrimidone) and T(dew)T(thymidylyl(3’→5’)thymidine) in vacuum, aqueous solution and also the (6-4) photolyase. In early experiments by Johns in 1964, upon excitation by UV light at 313 nm, the (6-4) photoproduct was found to convert into a new photoproduct which was confirmed to be the Dewar photoproduct.\textsuperscript{119} The reverse reaction happened upon excitation at the wavelength of 240 nm.\textsuperscript{119} However, to the best of our knowledge, the underlying mechanism of this photoisomerization reaction is still unclear to date. The optimized geometric structures for T(6-4)T and T(dew)T are shown in Figures 4.15 and 4.16 and corresponding energies are summarized in Table 4.2.

**Figure 4.15** Optimized geometries at CAS(10,8)/6-31G(d) level for the (6-4) PP in the ground states and excited states and also the conical intersections (CI). Selected bond lengths are shown in Å. Selected from paper IV, reprinted with permission from American Chemical Society.
Figure 4.16 Optimized geometries at CAS(10,8)/6-31G(d) level for the Dewar PP in the ground states and excited states and also the conical intersections (CI). Selected bond lengths are shown in Å. Selected from paper IV, reprinted with permission from American Chemical Society.

<table>
<thead>
<tr>
<th>6-4PP</th>
<th>CASPT2 (eV)</th>
<th>DewarPP</th>
<th>CASPT2 (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_0$</td>
<td>0.0</td>
<td>$S_0$</td>
<td>0.0</td>
</tr>
<tr>
<td>$T_1$</td>
<td>2.9</td>
<td>$T_1$</td>
<td>3.0</td>
</tr>
<tr>
<td>$S_1(S_n)$</td>
<td>3.2</td>
<td>$S_1$</td>
<td>4.2</td>
</tr>
<tr>
<td>$S_2(S_{pi})$</td>
<td>3.6</td>
<td>$S_2(S_n)$</td>
<td>4.6</td>
</tr>
<tr>
<td>$S_0S_2$</td>
<td>3.9</td>
<td>$S_0T_1$</td>
<td>4.1</td>
</tr>
<tr>
<td>$S_0S_1$</td>
<td>4.4</td>
<td>$S_nT_1$</td>
<td>5.1</td>
</tr>
<tr>
<td>$S_0T_1$</td>
<td>3.1</td>
<td>$S_0S_1$</td>
<td>5.4</td>
</tr>
<tr>
<td>$S_1T_1$</td>
<td>3.9</td>
<td>$T_1$-TS</td>
<td>3.1</td>
</tr>
<tr>
<td>$S_0$-TS</td>
<td>3.7</td>
<td>$S_n$-TS</td>
<td>5.6</td>
</tr>
</tbody>
</table>

Table 4.2 Calculated relative energies for (6-4) PP and Dewar PP in the ground and excited-states at CASPT2/6-31G(d) level based on the optimized structures at CASSCF(10,8)/6-31G(d) in gas phase. Selected from paper IV, reprinted with permission from American Chemical Society.
With this basic geometric and electronic structures, we draw a general picture for the photo-induced isomerization process between (6-4)PP and DewarPP in Figure 4.17.

![Figure 4.17](image)

**Figure 4.17** Calculated potential energy surfaces and suggested mechanism for the UV photoinduced isomerization between (6-4) PP and Dewar PP. The ground state is denoted by the black line. Moreover, the lowest triplet state T\(_1\) and singlet \(^1\pi\pi^*\) state S\(_n\) are shown in blue and green lines. Radical anion (RA) pathway is shown in magenta. The red and purple arrow lines show the possible isomerization mechanism between (6-4) PP and Dewar PP. Selected from paper IV, reprinted with permission from American Chemical Society.

As shown in Figure 4.17 with red lines, the (6-4) photoproduct is firstly excited to the S\(_{\pi}\) \((^1\pi\pi^*)\) state upon the UV excitation at 313nm. Along the potential energy surface of the S\(_{\pi}\) state, a nonradiative deactivation from the S\(_{\pi}\) state to the ground state takes place via the conical intersection —S\(_0\)/S\(_{\pi}\) which has a out-of-plane deformation shown in Figure 4.15. On the potential energy surface of the S\(_0\) state, the succedent bonding reaction via S\(_0\)-TS gives rise to the formation of the Dewar...
photoproduct. The conical intersection plays an important role in the photoisomerization.

The reverse process occurs in the \( S_n \left( ^1\pi\pi^* \right) \) state by means of the N3’-C6’ bond cleavage transition state \( S_n\text{-TS} \). However, in vacuum, the reaction barrier is too high to be overcome. The picture will change in the aqueous solution. We found two efficient isomeric pathways via the “dark” states \( S_n \) and \( T_1 \) in solution as discussed in section 4.1.

In addition, the magenta lines in Figure 4.17 represent the new electron-injection pathway which is found in (6-4) photolyase.\(^{108}\) Our DFT calculations indicate that the reaction barrier for \( S_0\text{-TS} \) is sharply decreased from 1.4 eV to 0.8 eV when one electron is injected to the neutral molecule. This radical anion way is suggested to be an efficient repair channel for the DewarPP, supporting the hypothesis from recent experiments.\(^{108}\)

4.2.3 DNA photorepair

In paper V, we explored the photophysical and photochemical properties of FADH\(^+\) which is the cofactor in DNA repairing process. Under the electron-transfer redox cycle of the excited FADH\(^+\) molecule, the photodamaged lesions are split through a series of repair reactions. In paper VI, we performed the density functional theory method to study several possible enzymatic repair pathways for the Dewar photoproduct—T(dew)C. The model introduced in section 4.1 has also been used in paper VI, see Figure 4.18. In order to investigate the special function of the two residues (His365 and His 369) for the DNA repair activity, we took His365 as neutral (HIE) and protonated (HIP) form in our calculations.

Several proposed mechanisms are depicted in Figure 4.19.
4.2 Computational photophysical and photochemical studies on UV-induced DNA issues

**Figure 4.18** Structure of the T(dew)C photoproduct (left) and the active site in (6-4) photolyase (right). PDB code: 2WQ6.

**Figure 4.19** Reaction mechanisms for the repair of T(dew)C.
The repair of the T(dew)C involves the transfer of the amino group from the C5 atom to C4’ atom and also the cleavage of N3’-C6’ bond. As depicted in Figure 4.19, there are concerted (path B) and step-by-step (path A and C) pathways for the amino group transfer. We studied these three paths under HIE365 and HIP365 options and the whole calculations were performed in the anionic ground state with the consideration of the electron donation from the FADH\(^-\). However, in both situations, the barriers of those amino group transfer pathways are very high; see red, blue and purple lines in Figure 4.20.

**Figure 4.20** Calculated potential energy profiles in different orientations with neutral and protonated His365.

It seems that the repair process will shut off through above channels. However, in the experiments of Glas et al.,\(^{108}\) they found the enzyme (6-4) photolyase was able to repair DewarPP into (6-4) lesions with one electron injection. Motivated by their experimental results, we then calculated the barrier of the N3’-C6’ cleavage that corresponded to the isomerization from T(dew)C into T(6-4)C in the anionic ground state. The optimized structures are shown in Figure 4.21 and the energies are drawn in green lines in Figure 4.20.
Figure 4.21 Optimized geometrical parameters with HIE 365 and HIP365 (in parentheses) at the UB3LYP/6-31G(d) level for radical anion in the photolyase.

The relative low barriers for this isomerization process indicate that the repair of DewarPP needs first isomerization to (6-4)PP in the (6-4) photolyase which is in good agreement with the experimental studies.\cite{108}
References


33. Coyle, J. D., Introduction to organic photochemistry, John Willey & Sons Ltd., 1986.


2010, 466, 887–890.


