A Triplet Mechanism for the Formation of Cyclobutane Pyrimidine Dimers in UV-Irradiated DNA

Ru Bo Zhang and Leif A. Eriksson*

Department of Natural Sciences and Örebro Life Science Center, Örebro University, 701 82 Örebro, Sweden

Received: January 11, 2006; In Final Form: February 21, 2006

The reaction pathways for the photochemical formation of cyclobutane thymine dimers in DNA are explored using hybrid density functional theory techniques. It is concluded that the thymine–thymine [2 + 2] cycloaddition displays favorable energy barriers and reaction energies in both the triplet and the singlet excited states. The stepwise cycloaddition in the triplet excited state involves the initial formation of a diradical followed by ring closure via singlet–triplet interaction. The triplet mechanism is thus completely different from the concerted singlet state cycloaddition processes. The key geometric features and electron spin densities are also discussed. Bulk solvation has a major effect by reducing the barriers and increasing the diradical stabilities. The present results provide a rationale for the faster cycloreaction observed in the singlet excited states than in the triplet excited states.

1. Introduction

Cyclobutane pyrimidine dimers (CPDs) are the major photoproducts observed in UV-irradiated DNA and have been correlated with cell lethality, mutagenesis, and the development of skin cancer. CPDs are formed between adjacent pyrimidine bases in DNA exposed to radiation in the far-UV (200–300 nm) range. The deleterious effects associated with these lesions largely result from the inhibition of DNA replication and transcription, and perturbation of interstrand base-pairing interactions subsequently leading to miscoding during DNA replication. To specifically recognize and repair these damages, organisms make use of a number of different enzymatic pathways. DNA photolyases, which are activated by concurrent UV absorption of nucleic acids initially produce localized triplet states than in the triplet excited states.

The stepwise cycloaddition in the triplet excited state involves the initial formation of a diradical from the concerted singlet state cycloaddition processes. The key geometric features and electron spin densities are also discussed. Bulk solvation has a major effect by reducing the barriers and increasing the diradical stabilities. The present results provide a rationale for the faster cycloreaction observed in the singlet excited states than in the triplet excited states.
are used to specify the state, and the number in superscript is a running index relating to the PES (e.g., TT$S_2$, TTT$S_5$, and so forth).

2. Computational Methods

Photoinduced dimerization reactions of large molecular aggregates are a major challenge for theoretical studies. The most appropriate approach in this context is to calculate the PES using a multireference approach such as complete active space second-order perturbation (CASPT2) theory to account for both dynamical electron correlation and near-degeneracy effects. Because of the computational requirements associated with such an approach, most quantum chemical investigations of photocycloadditions have, however, been carried out at the complete active space self-consistent field (CASSCF) level of theory and have primarily concerned photocycloadditions involving small alkenes or dienes. However, density functional theory (DFT)-based methods have made it possible to explore photocycloadditions in larger systems (such as current nucleoside dimers) using quantum chemical calculations, even though DFT in comparison with multiconfigurational ab initio theory in certain regards constitutes a less appropriate foundation for the modeling of photochemical reaction processes. The approximate density functionals used in current formulations of Kohn–Sham single-reference DFT may, for example, fail to describe systems in which near degeneracies are occurring between different electronic configurations. Our earlier time-dependent (TD)-DFT calculations yielded an approximate potential energy curve for the concerted singlet-state photochemical reaction pathway having a minimum (a “potential well”) at a nuclear configuration corresponding to that of the transition structure along the thermal reaction path, in accordance with the van der Lugt and Oosterhoff model of photochemical reactions.

In this paper, we have explored the possible triplet energy pathways to dimerization, assuming the system to undergo intersystem crossing after initial excitation to $S_1$ (or $S_2$). All triplet calculations were carried out using the B3LYP hybrid density functional as implemented in the Gaussian 03 suite of programs in conjunction with the 6-31G(d,p) and 6-311++G(d,p) basis sets. The computational model system is focused on the formation of the biologically relevant cis-→syn stereoisomer (cf. Scheme 1). The neglect of the influence of the DNA backbone (e.g., strain effects) and the complementary strand (e.g., hydrogen-bonded interactions) on the reacting bases is necessitated by the need to calculate numerous structures (see below) through which the systems pass on the route between reactants and products. Even though the surrounding DNA and the presence of a solvent most certainly affect the photodimerization, it may be noted that the absorption of light by DNA results in localization of energy primarily at the sites of the bases.

For the reactant complexes, unconstrained geometry optimizations lead to coplanar structures in which the bases interact by hydrogen bonds. As this is clearly an unphysical model of neighboring bases in DNA, we chose to optimize the reactant complexes with geometrical constraints (T + T, R(C5′→C5) = 4.18 Å, R(C6′→C6) = 4.46 Å, and C7′→C5′→C5′′ = 35.3°) to ensure a structure as in native DNA. No geometrical constraints were however enforced when optimizing transition structures or products. B3LYP/6-31G(d,p) frequency calculations were performed to identify the stationary points as either minima (reactant complexes and product structures) or first-order saddle points (transition structures) and to extract zero-point vibrational energy (ZPE) corrections. Bulk solvation effects were considered using the integral electron formalism of the polarized continuum model (IEF-PCM). In parallel, the calculations were also performed on the singlet states using the same geometric conformations. The correlated photochemical reaction pathways of the cycloadditions were explored on the basis of TD-DFT calculations developed by Scuseria and co-workers. This implementation computes vertical excitation energies only. In principle, this prevents the exact localization of stationary points on an excited-state potential energy surface. However, within the framework of the van der Lugt and Oosterhoff model in which the geometrical features of the excited-state path are similar to those of the ground-state path, an approximate excited-state potential energy surface can be calculated using ground-state geometries.

3. Results and Discussions

3.1. Cycloaddition on Singlet PES. First, the cycloaddition reactions on the singlet ground and excited states are briefly reviewed. The optimized singlet ground-state geometries of the initial complex (TT$S_1$), transition state (TT$S_2$), and cross-linked product (TT$S_3$), and the relevant geometric parameters, are shown in Figure 1.

For the initial complex (TT$S_3$), the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO), respectively, are mainly localized on each of the two bases, with an energy gap of 0.44 au. The stabilization energy of the complex TT$S_3$, relative to two isolated thymine bases, is 2.3 kcal mol$^{-1}$ at the B3LYP/6-31G(d,p) level. Along the concerted thermal path, the TT$S_2$ transition structure is associated with a barrier of 87.6 kcal mol$^{-1}$. The C5′→C5′ distance is 0.2 Å longer than the C6′→C6′ one, which is attributed to the repulsion between the two methyl groups. In the TT$S_2$ transition structure, puckering of the thymine moieties is furthermore observed, due to the conversion of the C5′→C6′→C5′′ dihedral angle is 18.1°. In the produced CPD, the individual C5′C6 bonds are typical to those of single bonds and with distances ca. 0.2 Å greater than those of the initial complex. The C5′→C6′→C5′′ dihedral angle is 19.9°. In comparison with the X-ray crystal structure of the TT dimer in deacetyDNA, given in parentheses of Figure 1, we note that the C5′C6 and C5′′C6′′ bonds in the crystal data are essentially still x-bonds (1.3–1.4 Å), whereas upon cycloaddition they should attain a complete $\sigma$-character. The low resolution (2.5 Å) of the X-ray data may be a possible explanation for this. The bridging C5′→C5′ and C6′→C6′′ distances are in good agreement between experiment and theory.
Figure 1. Selected UB3LYP/6-31G(d,p) geometrical parameters of the stationary structures along the thymine–thymine cycloaddition reaction in their singlet and lowest-lying triplet states. For the singlet CPD product TT$_S^3$ X-ray data from PDB entry 1N4E$^{10}$ is added in parentheses.
singlet route to the formation of CPDs has recently been verified to the ground-state surface. As mentioned above, the excited singlet state; (b) triplet state.

The energy of TTT\(^1\) used as zero in the table is -908.301467 au (for vacuum) and -908.330264 au (for \(\epsilon = 4.3\)) at the B3LYP/6-31G(d,p) level and -908.551203 au (vacuum) and -908.569890 au (\(\epsilon = 4.3\)) at the B3LYP/6-311++G(d,p) level.

The potential energy curves of the thermal and photochemical singlet reaction pathways are presented in Figure 2a; relative energies are given in Table 1. The total reaction energy is 18.0 kcal mol\(^{-1}\). Both the high barrier and endothermicity illustrate that thermally induced CPD formation is energetically inaccessible. The singlet photochemical process (S\(_1\) surface) of CPD formation reveals an energy barrier of only 3–4 kcal mol\(^{-1}\) before the system reaches the potential well from which it decays to the ground-state surface. As mentioned above, the excited singlet route to the formation of CPDs has recently been verified by time-resolved experiments.\(^{23}\)

3.2. Cyclodaddition on the Lowest-Lying Triplet Excited PES. As noted in earlier studies,\(^{18,40}\) the main structural feature of the lowest-lying triplet of pyrimidine bases is the conversion of the C5–C6 \(\pi\)-bond (1.352 Å) to a \(\sigma\)-bond (1.501 Å). For the isolated thymine base, the singlet–triplet adiabatic energy gap is 65.0 kcal mol\(^{-1}\) at the B3LYP/6-31G(d,p) level. The spin densities are localized on the reactive C5 and C6 atoms. The electronic features imply that the cyclodaddition mechanism on the triplet PES should be different from the one on the singlet surface.

The first excited triplet, TTT\(^1\), can be reached via ISC from the excited singlet surface. The optimized lowest-lying triplet stationary points are shown in Figure 1. In the initial triplet complex TTT\(^1\), the CSC6 distance of one fragment of TTT\(^1\) is the same as for the singlet thymine base, showing that it retains the properties of the singlet ground state. The CSC6 distance of the other unit is 1.488 Å, which is close to the length formed in the triplet thymine base. A dramatic puckering structure is observed for this fragment, and the H(C6′) deviates considerably from the quasi-plane. The unpaired spin densities are mainly localized on C6′ (0.82 electrons) and C5′ (0.75 electrons), as seen in Table 2. The stabilization energy of the TTT\(^1\) complex is 4.2 kcal mol\(^{-1}\), relative to the energetic sum of isolated triplet and singlet bases, which is slightly higher than the stabilization energy of the singlet ground-state complex TTS\(^1\) compared to two singlet molecules. The adiabatic singlet–triplet energy gap between TTS\(^1\) and TTT\(^1\) is 65.7 kcal mol\(^{-1}\), which is very close to the single base case.

The reaction energy profiles of triplet cycloaddition reactions are presented in Figure 2b. One of the paths to the formation of cis–syn CPDs on the triplet PES proceeds through initial attack by C5 on C5′, followed by C6–C6′ cross-link formation (route TTT\(^1\)–TTT\(^2\)–TTT\(^4\)–TTT\(^6\)–TTT\(^8\)). From TTT\(^1\) to TTT\(^2\), the activation energy is 11.3 kcal mol\(^{-1}\) at the B3LYP/6-31G(d,p) level (10.3 kcal mol\(^{-1}\) at the B3LYP/6-31+1G(d,p) level). In the transition structure the base puckering is lost, and repulsion between the two methyl groups makes these deviate from the individual base planes. The C5′–C5′ distance is 2.251 Å, which is slightly shorter than that in TTS\(^2\). The C6′–C6′ distance is 2.834 Å, and the two unpaired electrons are now more delocalized over both fragments, indicative of spin transfer from one base to the other. The \(\Delta E_{C5′–C6′–C5′}\) dihedral angle is 27.1°; i.e., the two bases are now ca. 8° less rotated relative to each other, compared to the situation in native DNA (dihedral angle 35.3°). The resulting C5–C5′ cross-linked intermediate TTT\(^3\) is a diradical with spin densities localized to C6 and C6′ (0.84 e\(^{-}\) on each). The \(\Delta E_{C5–C6–C5′}\) dihedral angle is now 33.9°. The reaction energy relative to TTT\(^1\) is +2.8 (+3.5) kcal mol\(^{-1}\) at the B3LYP/6-31G(d,p) (B3LYP/6-311++G(d,p)) level. The energetic difference between TTT\(^2\) and TTT\(^4\) is approximately 8.5 (6.8) kcal mol\(^{-1}\).

The reaction energy from the intermediate TTT\(^3\) to the transition structure for ring closure on the triplet surface, TTT\(^4\), 49.4 kcal mol\(^{-1}\) must be

---

**TABLE 1: Relative Energies\(^a\) (without ZPE Corrections), in kcal mol\(^{-1}\), of the Species along the Thymine–Thymine Cyclodaddition Reactions**

<table>
<thead>
<tr>
<th>B3LYP/6-31G(d,p)</th>
<th>B3LYP/6-31G(d,p)</th>
<th>B3LYP/6-311++G(d,p)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(vaccum)</td>
<td>((\epsilon = 4.3))</td>
</tr>
<tr>
<td>TTS(^1)</td>
<td>0.0</td>
<td>65.7</td>
</tr>
<tr>
<td>TTS(^2)</td>
<td>76.6</td>
<td>75.5</td>
</tr>
<tr>
<td>TTS(^3)</td>
<td>72.5</td>
<td>70.5</td>
</tr>
<tr>
<td>TTT(^1)</td>
<td>68.5</td>
<td>68.4</td>
</tr>
<tr>
<td>TTT(^2)</td>
<td>56.7</td>
<td>56.6</td>
</tr>
<tr>
<td>TTT(^3)</td>
<td>117.9</td>
<td>114.8</td>
</tr>
<tr>
<td>TTT(^4)</td>
<td>115.1</td>
<td>115.1</td>
</tr>
<tr>
<td>TTT(^5)</td>
<td>109.9</td>
<td>110.3</td>
</tr>
</tbody>
</table>

\(^a\) The energy of TTS\(^1\) used as zero in the table is -908.301467 au (for vacuum) and -908.330264 au (for \(\epsilon = 4.3\)) at the B3LYP/6-31G(d,p) level and -908.551203 au (vacuum) and -908.569890 au (\(\epsilon = 4.3\)) at the B3LYP/6-311++G(d,p) level.
added at the B3LYP/6-31G(d,p) level. The result is little influenced by additional polarization or diffuse basis functions. The H atoms attached to C6 and C6′ in the transition structure TTf6 deviate from the individual plane with dihedral angles \( \angle H6-N1-C5-C6 = -23.2^\circ \) and \( \angle H6'-N1'-C5'-C6' = -24.9^\circ \). The \( \angle C5-C6-C6'-C5' \) dihedral angle is reduced to 14.8°, and the density is mainly located on C4−O4 of one base. The energetic difference between TTf6 and TTf5, the final CPD triplet-state product, is \( -8.0 \) (7.5) kcal mol\(^{-1} \). For TTf8, the C5C′ and C6C′ bond lengths are 1.615 and 1.581 Å, respectively, slightly longer than the corresponding distances in the singlet ground-state cycloduct TTt5, and the C5C6 and C5′C6′ bond lengths are 1.563 and 1.550 Å, respectively, which are almost consistent with the ones in TTt5. The unpaired spin electrons in TTf8 are localized on C4 (0.67 e\(^{-}\)) and O4 (0.99 e\(^{-}\)) of one base, similar to the case of TTf6. It is noteworthy that the singlet−triplet adiabatic energy gap between TTt3 and TTf8 is 92.0 kcal mol\(^{-1} \) at the B3LYP/6-31G(d,p) level, i.e., considerably higher than for the initial reactants TTt5 and TTf4.

The other pathway leading to the formation of cis-syn CPD on the triplet PES is by initial C6 attack to C6′ followed by subsequent formation of the C5−C5′ cross-link. Along the TTf1−TTt3−TTt5−TTt7−TTt8 route, an initial barrier of 6.8 kcal mol\(^{-1} \) must be overcome at the B3LYP/6-31G(d,p) level (5.3 kcal mol\(^{-1} \)) at the B3LYP/6-311++G(d,p) level) to reach TTt3, which is lower than the one leading from TTt1 to TTt4. In the transition structure TTt3, the C6−C6′ and C5−C5′ distances are 2.466 and 3.001 Å, respectively, and the \( \angle C5-C6-C6'-C5' \) dihedral angle is 27.2°. Spin densities are mainly found on C5 and C6′ but are starting to delocalize over both moieties. The TTt5 intermediate is also a diradical with a C6C′ bond length of 1.602 Å, which is slightly shorter than the C5−C5′ distance in TTt4. The C5−C5′ distance is 3.120 Å and the \( \angle C5-C6-C6'-C5' \) dihedral angle is 60.8°. Spin densities are, as expected, located on C5 and C5′ (ca. 0.84 e\(^{-}\)) on each. The TTt7 diradical is more stable than TTt4 by 11.8 (12.3) kcal mol\(^{-1} \). The reaction energy for TTt4−TTt6 is −9.0 kcal mol\(^{-1} \) at the B3LYP/6-31G(d,p) level. Thus, the initial cross-linking reaction TTf4−TTt6 is preferred over TTt4 since the C6−C6′ bond is preferential to the one involving the C5−C5′ bond, primarily due to reduced steric repulsion between the C5 methyl groups in TTt3 compared with TTt2. From TTt5 to TTt7, the activation energy is 58.5 (68.5) kcal mol\(^{-1} \), which is comparable to the barrier for the reaction TTt4−TTt6. For the transition structure TTt7, the C5−C5′ distance is 1.904 Å, and C6−C6′ is nearly unchanged compared to the distance in TTt5. The \( \angle C5-C6-C6'-C5' \) dihedral angle is 20.6°. The energy difference between TTt7 and TTt8 is −5.2 (−4.6) kcal mol\(^{-1} \).

According to the above results, the T'T' dimer cannot be formed on the triplet PES only, due to the high barriers for both the TTt4−TTt6−TTt8 and TTt5−TTt7−TTt8 reactions. The present results hence do not agree fully with the results implied by various experiments that the T'T' dimer can potentially be formed through the triplet excited state alone. Instead, knowledge of the interaction between the lowest-lying triplet and the singlet ground-state surface is important for an understanding of the triplet CPD formation.

To obtain a reliable singlet−triplet interaction profile, the PES was scanned from TTt4 to TTt6 by varying the C6−C6′ distances with a step length of 0.1 Å, constraining the C5−C5′ distance at 1.644 Å and optimizing the remaining coordinates. The same procedure was used to determine the PES from TTt5 to TTt7, fixing C6−C6′ at 1.602 Å and scanning the C5−C5′ distance.

The energies for the singlet/triplet scans are displayed in Figure 3a (starting from TTt4) and 3b (starting from TTt5). For system TTt4, with C5−C5′ already formed, there is a triplet−singlet crossing point at the C6−C6′ distance of 2.335 Å with an energy of ca. −908.169952 au at the B3LYP/6-31G(d,p) level. Thus the estimated energy gap from TTt4 to the crossing point is approximately 14.1 kcal mol\(^{-1} \). For the path TTt5 to TTt7 (Figure 3b), the C5−C5′ distance at the crossing point is close to 2.397 Å, and the energy gap from the initial TTt5 complex to the crossing point is estimated to be 14.9 kcal mol\(^{-1} \). The determined energy barriers to the spin crossover points are dramatically lower than the 87.6 kcal mol\(^{-1} \) barrier for the thermal cyclodaddition reaction and the ca. 50 kcal mol\(^{-1} \) required to reach the triplet CPD, indicating that the triplet state is a possible channel to the formation of the T'T' dimer if interaction with the singlet surface is considered after formation of the initial single cross-linked diradical intermediate.

### 3.3. Effects of Bulk Solvation

In this section, we will address the role of bulk solvation on the reaction profiles. The dielectric constant \( \varepsilon = 4.3 \) is used to simulate the local surroundings of DNA. In Table 1, the energetics derived from the density functional theory self-consistent reaction field (DFT-SCRF) calculations are listed. Several features can be observed from the reaction profiles after accounting for the implicit solvent effect generated by the nonaqueous medium. Relative to the gas-phase value, the stability of TTt3 is enhanced by 18.1 kcal mol\(^{-1} \) due to bulk solvation effects. The barrier heights decrease in the TTt4−TTt6 and TTt5−TTt7 reaction steps to 6.0 and 1.1 kcal mol\(^{-1} \), respectively, compared to the corresponding gas-phase values of 11.3 and 6.8 kcal mol\(^{-1} \) at the B3LYP/6-31G(d,p) level. Essentially the same result is obtained when the larger basis set is considered. The reaction energies for the formation of the diradicals, for example, TTt4−TTt5 and TTt4−TTt1, are −1.1 and −12.9 kcal mol\(^{-1} \) at the B3LYP/6-31G(d,p) level (gas-phase values of +2.8 and −9.0 kcal mol\(^{-1} \), respectively). The bulk effects of the surroundings even at such a low value of \( \varepsilon \) as 4.3 D are hence considerable, some 4−6 kcal/mol on both reaction rates (barrier heights) and stabilities of the two diradical intermediates.

DFT-SCRF is also used to study the bulk solvent effects on the interaction between singlet and triplet states. As seen in Figure 3, the sites of the crossing points are essentially unchanged by the influence of bulk solvent, although their stabilities are enhanced. The energy gaps from the crossing points to the initial TTt4 and TTt5 complexes are estimated to 13.2 and 14.3 kcal mol\(^{-1} \), respectively, which are slightly lower than the corresponding gas-phase values.

### 4. Conclusions

In the present work, hybrid DFT methods have been employed to investigate the processes of CPD formation on the singlet and lowest-lying triplet excited states with the objective to determine reaction energy profiles. Geometries, unpaired electron spin densities, and reaction energies were obtained at the B3LYP/6-31G(d,p) level in the gas phase, followed by energy calculations performed at the same level in solution (\( \varepsilon = 4.3 \)) using the IEF-PCM model. Single-point calculations were also performed, both in vacuo and in bulk solvent, using the larger 6-311++G(d,p) basis set. Two stepwise pathways to CPD formation are found on the triplet PES, and the results are compared to the concerted mechanism of formation on the
singlet surface. For the initial reactions (formation of the first C—C' cross-link bond, either C5—C5' or C6—C6') on the triplet PES in gas phase, the barriers for the two paths are between 6 and 11.5 kcal mol\(^{-1}\) and result in the formation of diradical intermediates. The energetically most favorable path, both in terms of activation energy and diradical stability, is that initiated by formation of the C6—C6' bond (TT\(_4^5\) intermediate), as compared with the formation of TT\(_4^4\) (the C5—C5' cross-linked intermediate) in which the methyl groups on C5 and C5' present unfavorable interactions.

The reactions from the diradicals to the ring-closed CPDs are prohibited by high energetic barriers, comparable to the one of the thermal singlet ground-state process. Instead, the interaction between triplet and singlet states close to the diradical intermediates plays a significant role in the CPD formation. Our estimated energy differences from the two diradicals to the corresponding singlet—triplet crossing points range from 14 to 15 kcal mol\(^{-1}\) and may pose a possible channel for CPD formation via the triplet PES. Bulk solvation effects lead to a dramatic decrease of the initial energetic barriers and reaction energies, whereas the energy gaps between the diradicals and the singlet—triplet crossing points are little affected.

The current data provide a rationale for the observation that triplet mechanisms may be involved in CPD formation, albeit at a slower rate than on the excited singlet surfaces. A crucial factor in this respect is the possibility for the initially excited singlet system to undergo intersystem crossing to reach the triplet complex TT\(_1^4\).

Acknowledgment. The Swedish Science Research Council (VR) is gratefully acknowledged for financial support. We also acknowledge generous grants of computing time at the National supercomputing facilities in Linköping (NSC) and Stockholm (PDC).
References and Notes