Femtosecond Stimulated Raman Spectroscopy of the Cyclobutane Thymine Dimer Repair Mechanism: A Computational Study

Hideo Ando,† Benjamin P. Fingerhut,‡ Konstantin E. Dorfman,† Jason D. Biggs,† and Shaul Mukamel*†

†Department of Chemistry, 1102 Natural Sciences II, University of California, Irvine, California 92697-2025, United States
‡Max-Born-Institut für Nichtlineare Optik und Kurzzeitenspektroskopie, Max-Born-Strasse 2A, 12489 Berlin, Germany

Supporting Information

ABSTRACT: Cyclobutane thymine dimer, one of the major lesions in DNA formed by exposure to UV sunlight, is repaired in a photoreactivation process, which is essential to maintain life. The molecular mechanism of the central step, i.e., intradimer C–C bond splitting, still remains an open question. In a simulation study, we demonstrate how the microscopic photoreactivation mechanism shows only absorptive features in the simulated FSRS signals, whereas the fast concerted mechanism shows characteristic dispersive line shapes.

1. INTRODUCTION

The photoinduced excited-state dynamics of DNA is of great importance in biology, medicine, and life science. The exposure of living organisms to UV sunlight causes harmful lesions to DNA. One of the major lesions is cyclobutane pyrimidine dimer (CPD), which is formed through [2+2] photocycloaddition and may eventually lead to skin cancer. Living organisms often use specific flavoproteins, CPD-photolyase (PL), to repair the lesions in DNA sequences using a blue light activated enzymatic cycle (see Scheme 1). Macroscopically, this photoreactivation process satisfies Michaelis–Menten kinetics, where CPD-containing DNA is bound to the redox cofactor flavin-adenine dinucleotide (FADH+) in CPD-PL and light acts as splitting agent. Recent ultrafast time-resolved transient absorption experiments and simulations revealed that the microscopic photoreactivation mechanism involves three steps as summarized in Scheme 1: electron transfer from excited FADH+ to CPD, splitting of intradimer C–C bonds in CPD, and electron return to restore catalytically active FADH+, whose rates are well optimized to ensure a high repair quantum yield (about 0.5–1.0). The entire photoreactivation process is completed within one nanosecond. For the second and central step of the intradimer C–C bond cleavage, a number of transient absorption studies have been carried out to clarify the kinetics. In the overall splitting process of the two C–C bonds, however, there is controversy due to the lack of unique molecular probes. Barrierless formation of one C5=C6 double bond upon electron transfer was observed within 60 ps by MacFarlane and Stanley and later associated with overall C–C splitting by Masson et al. and Langenbacher et al. proposed slower splitting rates (i.e., 260 ps, <560 ps, and no splitting below 200 K), and an overall barrier was estimated to be 10.7 ± 2.3 kcal/mol. Besides, transient absorption results by Liu et al. suggested that CPD splitting occurs in a sequential mechanism, where the two bonds are broken in a stepwise manner: 10 ps splitting of the C5=C5 bond, followed by 90 ps splitting of the C6=C6 bond (see Scheme 1). In contrast, theoretical investigations of CPD embedded in the active site of CPD-PL and in DNA duplex suggest ultrafast CPD splitting in an (asynchronous) concerted mechanism, whereby the C5=C5′ bond splits upon electron uptake within 10–100 fs and the C6=C6′ bond splits within a few picoseconds. Interpretation of visible-light probe studies is difficult because relevant radical reaction intermediates as well as other radicals arising from photoreactivation processes absorb in this region. On the other hand, UV probes monitor the appearance of repaired thymine, but provide only limited sensitivity for transient radical intermediates. Details of the splitting mechanism remain an open question and a direct spectroscopic probe of the molecular rearrangements upon the CPD splitting pathway is yet to be reported.

Time- and frequency-resolved vibrational spectroscopy with infrared or Raman probes can closely follow specific atomic motions. Unique vibrational bands serve as fingerprints of the molecular rearrangements.
excited-state photochemistry and photophysics, and their real-time observation can show transient reaction intermediates\(^1\)\(^-\)\(^3\)\(^1\) and reveal reaction mechanisms.\(^2\)\(^2\)\(^3\) In a UV/visible pump \(\rightarrow\) Raman probe experiment, a pump pulse excites the molecular system into a valence excited state, and a delayed Raman sequence probes the subsequent rearrangement process. In femtosecond stimulated Raman spectroscopy (FSRS),\(^2\)\(^4\) a Raman probe sequence consists of an off-resonant picosecond probe \(k_2\) and a superimposed femtosecond laser pulse \(k_3\), which stimulates the Raman signal, and successive spectra can be recorded with \(\Delta T \approx 20\) fs time intervals with high spectral resolution. Following the original work of Yoshizawa and Kurosawa,\(^2\)\(^4\) this technique provides a sensitive local probe for ultrafast light-induced processes.\(^2\)\(^1\)\(^2\)\(^3\) Different configurations of the FSRS techniques including temporally and spectrally overlapping pulses and resonant Raman processes,\(^2\)\(^6\) and cascading effects in FSRS\(^2\)\(^7\) have been calculated.\(^2\)\(^8\)

Recently we developed an intuitive picture of FSRS signals based on a loop diagram representation.\(^2\)\(^9\)\(^3\)\(^0\) The relevant molecular response is expressed by a multipoint correlation function, which can be obtained by microscopic quantum simulations. We have shown that even though the delay time \(T\) and spectral resolution are independent experimental knobs, the effective temporal and spectral resolution of the technique

\(\text{Scheme 1. Proposed Photoreactivation Mechanism of the CPD Lesion by CPD-PL}^a\)

\(\text{Figure 1. Optimized geometry of TpT}^{\bullet \bullet} \text{ in the closed form: (a) front and (b) side views. The TT}^{\bullet \bullet} \text{ moiety is marked by red dashed circle.}\)
is affected by the probed system dynamics and is inherently limited by the Fourier uncertainty $\Delta \omega \Delta t > 1$,21,31 are not to be interpreted as instantaneous snapshots of the nuclear frequencies29 and their interpretation requires a careful analysis.32,33

In this paper, we apply this theoretical approach to study how FSRS can monitor the repair dynamics of the cyclobutane thymine dimer radical anion (TT$^\ast$). The potential energy calculated at the density functional theory (DFT) level involves two minor ($<1$ kcal/mol) activation barriers for the C5–C6 splitting and a main barrier (5.4 kcal/mol) for the C6–C6 splitting. Spectroscopic signatures of transient intermediates during the bond rearrangements are identified. The high temporal resolution of the technique allows one to pinpoint the reaction mechanism upon electron uptake by directly monitoring the evolution of characteristic marker bands (i.e., C==O and C==C/C==C stretch vibrations). We find that the simulated FSRS signals for the sequential mechanism show absorptive peaks ($\sim$1630 cm$^{-1}$) of two C==C stretch vibrations due to the C6–C6 splitting, following the C5–C5 splitting. The concerted mechanism, in contrast, yields a dispersive line shape ($\sim$1560 cm$^{-1}$). We further investigate the repair pathway of CPD dinucleotide thymidylyl(3′→5′)thymidine radical anion (TpT$^\ast$), which is a first model step toward the CPD lesion embedded in the DNA strand (Figure 1).

2. CPD SPLITTING IN THE THYMINE DIMER RADICAL ANION

In the following, we investigate how FSRS can distinguish between the proposed sequential and concerted splitting mechanisms. We first study TT$^\ast$ which is a simple model for TpT$^\ast$ (see the circled moiety of the TpT$^\ast$ in Figure 1a). Here we focus on the intradimer splitting mechanism subsequent to the forward electron transfer (Scheme 1) as to be discussed in section 2.4, and ignore the CPD-PL enzyme. A potential energy curve of the intradimer splitting is constructed, and the evolution of characteristic vibrational marker bands is identified for the simulation of FSRS signals.

2.1. Optimized Geometries and the Intrinsic Reaction Coordinate (IRC) Path for the Intradimer Splitting Process. The TT$^\ast$ model has two intradimer bonds: the C5–C5 and C6–C6 bonds shown in Figure 1a. Geometry optimization yielded three local minima, namely the closed (dimerized form), the open form,34 and one intermediate state (INT). The two transition states (TSs) between the local minima were then obtained with the B3LYP algorithm. The reaction pathways from the TSs to local minima were explored with the intrinsic reaction coordinate (IRC) method,35 which gives the steepest descent pathway in mass-weighted Cartesian coordinates and yielded another INT and TS. Altogether there are two INTs and three TSs along the complete path toward the splitting (Figure 2). To study the CPD splitting mechanism of the closed form, we constructed a one-dimensional reaction coordinate via the TSs and INTs by connecting the resulting IRCs.36 Stability of the electronic state along the reaction coordinate was confirmed by the self-consistent-field stability analysis.37,38 Frequency analysis showed that TSs have a single imaginary frequency mode whereas the other geometries have only real-valued frequencies. All quantum chemical calculations employed the 6-31G* basis set39,40 and the unrestricted DFT(M05-2X) method41 to consider the open-shell doublet state of TT$^\ast$. The M05-2X is a hybrid meta exchange-corr

Table 1. Potential Energy Barriers (kcal/mol), Compared to the Preceding Stationary Geometries, in the TT, TT$^\ast$, and TpT$^\ast$

|     | TT   | TT$^\ast$ | TpT$^\ast$ |
|-----|------|----------|-----------|
| TS1 | 49.9 | 0.2      | 0.2       |
| TS2 | 1.5  | 0.9      | 0.5       |
| TS3 | 5.4  | 4.9, 5.2 | 4.4       |

$^a$In parentheses are the energies obtained with the CCSD and the CCSD(T) method at the DFT(M05-2X) optimized geometries.

In a DFT study of an analogue of the TpT$^\ast$, the barrier (with zero-point energy corrections) for the C5–C5 splitting is 1.8 kcal/mol and that for the C6–C6 splitting is 3.2 kcal/mol.43 In QM/MM dynamics studies of CPD embedded in the active site of CPD-PL,18 and in DNA duplexes,44 the C5–C5 splitting is barrierless and the free energy barrier of the C6–C6 splitting is less than 2.5 kcal/mol.
The potential energy along the IRC path of CPD splitting is depicted in Figure 2. As shown in Table 1, the potential energy barriers of TS1•− and TS2•− are less than 1 kcal/mol, while the last TS3•− barrier is the highest (5.4 kcal/mol). In Figure 3a, the intradimer C5—C5′ and C6—C6′ bond lengths are plotted along the IRC path. At INT1•, the C5—C5′ bond is partially cleaved (2.48 Å) and the C6—C6′ bond is intact (1.56 Å). As the geometry is relaxed to INT2•, the C5—C5′ bond fully splits (3.22 Å) but the C6—C6′ bond remains intact (1.67 Å). In addition, the C5—C6—C6′—C5′ dihedral angle opens by 35.2° and the displacement sum over all atoms in two thymine bases, compared to the free thymine structure, is decreased by 2.44 Å. These indicate that the small TS1•− and TS2•− barriers correspond to the C5—C5′ splitting and the TS2•− additionally involves an internal rotation around the C6—C6′ bond and relaxation of thymine structures themselves. Previous quantum chemical calculations had found only one (1.8 kcal/mol) or no TS for the C5—C5′ splitting, probably due to the tiny barriers and the flat potential energy curve. At TS3•−, both intradimer bond lengths are larger than 3.0 Å and the main barrier of 5.4 kcal/mol is attributed to the C6—C6′ splitting. This barrier height is in reasonable agreement with a previous study on the level of B3LYP/6-311++G(2df,p) theory (2.3 kcal/mol). The geometric characteristics are labeled using compact notation in the caption of Figure 2 (e.g., (T—̈T)•− for INT1•− and (T—̈T)•− for INT2•−).

Figure 2 additionally depicts the IRC potential energy of neutral thymine dimer (TT) splitting. It shows a highly activated stepwise mechanism, where TS1 and TS2 are attributed to the C5—C5′ and the C6—C6′ splitting. Key geometric parameters of INT are similar to those of INT2•, (T—̈T)•−, as shown in Figure S6 in Supporting Information. The 49.9 kcal/mol TS1 barrier may not be overcome by thermal activation, demonstrating the catalytic function of electron uptake in the radical anionic splitting pathway of TT•−; here only small barriers appear and the C5—C5′ splitting proceeds exothermically. This result agrees well with the experimental findings that electron uptake by the thymine dimer is necessary for both the CPD-PL catalyzed photoreactivation and the spontaneous self-repair.

2.2. Vibrational Marker Bands along the IRC Path. We selected six normal modes (modes 67–72) as characteristic marker bands: four C=C and the C=O stretch vibrations. Filled circles represent the optimized geometries.
that these frequencies are sensitive to bond rearrangements along the IRC path as shown in Figure 3b. In addition, the 1600−1900 cm\(^{-1}\) \(\text{C}==\text{O}\) and \(\text{C}==\text{C}\) stretch frequencies (see Table S2 for a thymine base) are well separated from those of the other modes, and show relatively strong Raman activity as seen in Figure S7.

In the 1600−1900 cm\(^{-1}\) frequency range (Figure 3b), four normal modes appear with a displacement of 0.0 (i.e., the closed form) to about 60 Bohr amu\(^{-1}\), which substantially change along the reaction coordinate, leading to a transient appearance of six modes up to 74.9 Bohr amu\(^{-1}\) (i.e., TS3\(^{-}\)) and five modes around the open form. These three regions are hereafter denoted by I, II, and III (see Figure 3b). Considering the geometric characteristics in section 2.1 and Figure 3a, the geometry change in region I corresponds to the \(\text{C}_5−\text{C}_4\) splitting combined with internal rotation around the \(\text{C}_6−\text{C}_5\) bond, while the transition from region I to II initiates the \(\text{C}_6−\text{C}_5\) splitting and that from region II to III finalizes the splitting.

The frequency variations in each region can be rationalized by the evolution of relevant molecular orbitals, particularly the singly occupied molecular orbitals (SOMOs); see Figures 4 and S8 for the singly occupied natural orbitals. The closed form has saturated \(\text{C}_4−\text{C}_3\) and \(\text{C}_5−\text{C}_6\) bonds, but no \(\text{C}_4==\text{C}_5\) double bonds. This is clear from the doubly occupied bonding orbital built from the lowest unoccupied molecular orbitals (LUMOs) of both thymine bases, which has antibonding \(\pi^*(\text{C}_4==\text{C}_5)\) character (see the left orbital diagram in Figure S5). The closed form only has four \(\text{C}==\text{O}\) stretch modes in the 1600−1900 cm\(^{-1}\) frequency range. The low 1615 cm\(^{-1}\) frequency of one of the \(\text{C}_4==\text{O}\) modes, localized in the left thymine in Figure 4, reflects the notable antibonding \(\pi^*(\text{C}_4==\text{O})\) character of the SOMO; however, as the \(\text{C}_4==\text{C}_3\) bond partially splits, the antibonding character decreases, which strengthens the \(\text{C}_4==\text{O}\) bond and leads to a blue shift of the frequency (cyan line in Figure 3b).

In the other parts of region I, which mainly correspond to the internal rotation around the \(\text{C}_6−\text{C}_5\) bond, the four \(\text{C}==\text{O}\) stretch modes remain within the 1600−1900 cm\(^{-1}\) range and their frequencies do not vary significantly except around TS2\(^{-}\). As the molecular geometry gets closer to INT2\(^{+}\), the SOMO gradually changes from the antibonding interaction between the highest occupied molecular orbitals (HOMOs) of the thymine bases, into the bonding interaction between the LUMOs, as can be seen in Figures 4 and S9. Although the SOMO changes are large, the local electron distribution around the four \(\text{C}==\text{O}\) bonds does not vary significantly and the frequency variations are thus small. Note that only when the geometry passes TS2\(^{-}\), does the SOMO become localized predominantly on one \(\text{C}_4==\text{O}\) bond and the \(\text{C}_5==\text{C}_6\) bond (see the sixth panel in Figure S9), resulting in a blue frequency shift of the other \(\text{C}_4==\text{O}\) stretch. The Duschinsky rotation matrices\(^{52}\) were calculated to investigate the mixing of normal modes accompanied by geometry change along the reaction coordinate. Before and after the boundary of regions I and II (see Figure S10), elements of the Duschinsky rotation submatrix of the lower two frequency modes are small, as seen in the lower left panel in Figure S11. In region II, the two normal modes thus completely change into \(\text{C}_5−\text{C}_4\) stretch due to the \(\text{C}_6−\text{C}_5\) splitting, and we find a total of six modes in the 1600−1900 cm\(^{-1}\) regime. The ~1630 cm\(^{-1}\) \(\text{C}_5==\text{C}_6\) stretching frequencies are lower than common \(\text{C}==\text{C}\) values (e.g., 1774 cm\(^{-1}\) in Table S2). This is because up to TS3\(^{-}\), or \((\text{T}−\text{T})\)\(^{−}\), the SOMO is built from both LUMOs of thymine bases and has strong antibonding \(\pi^*(\text{C}_5==\text{C}_6)\) character (Figure 4). Before and after the boundary of regions II and III (Figure S10), the Duschinsky submatrix of the six modes has large off-diagonal elements (see the lower middle panel in Figure S11).

As the geometry passes TS3\(^{-}\) and the \(\text{C}_6==\text{C}_5\) splitting is completed, one \(\text{C}_4==\text{C}_5\) stretch vibration, localized in the left thymine in Figure 4, disappears and we only find five modes in the range 1600−1900 cm\(^{-1}\). This is because the SOMO becomes localized in the left thymine; the system can be viewed as consisting of a neutral and a radical anion thymine base, which has five normal modes in the range 1600−1900 cm\(^{-1}\) as given in Table S2.

### 2.3. Two Model Trajectories for the Concerted and the Sequential Mechanism.

We considered an exponential model for the time-dependent IRC:

\[
R(t) = \begin{cases} 
R^{\text{open}} + e^{-\tau/T_0} (R^{\text{closed}} - R^{\text{open}}) & \tau \geq 0 \\
R^{\text{closed}} & \tau < 0 
\end{cases}
\]

where \(R^{\text{open}}\) and \(R^{\text{open}}\) are the IRC coordinate values of the closed and open forms and \(T_0\) is the geometry change time scale. This equation together with the DFT vibrational frequencies \(\omega_m(R)\) for a given mode along the IRC path gives a frequency trajectory \(\omega_m(R(t))\).

The rate of the \(\text{C}_4==\text{C}_5\) splitting depends on how fast the molecular geometry passes through region I (region II). More importantly, the lifetime of INT2\(^{+}\), or (\(\text{T}−\text{T}\))\(^{+}\), is crucial for distinguishing between the sequential\(^6\) and the concerted nature\(^{14,18}\) of the intradimer splitting mechanism. In QM/MM simulations,\(^{18}\) Masson et al. demonstrated that the \(\text{C}_5==\text{C}_4\) bond partially splits within 10−100 fs upon electron uptake. Both intradimer bonds then fully split within 1 ps. After the \(\text{C}_5==\text{C}_4\) splitting, the distance fluctuates around 2.70 ± 0.14 Å with a widespread range from 2.3 to 3.0 Å, and the \(\text{C}_5==\text{C}_4\)−\(\text{C}_5==\text{C}_4\) angle fluctuates around 35° ± 4° with a spread of up to 55°. This indicates that the fluctuating geometries are still far from INT2\(^{+}\), whose \(\text{C}_5==\text{C}_4\)−\(\text{C}_5==\text{C}_4\) distance and \(\text{C}_5==\text{C}_4\)−\(\text{C}_5==\text{C}_4\) angle are 3.22 Å and 67.5°, respectively. The second splitting process starts before the completion of the internal \(\text{C}_6==\text{C}_5\) rotation.

We use two trajectories: (A) at \(t_0 = 500\) fs and (B) at \(t_0 = 100\) ps. Trajectory A represents the concerted mechanism mentioned above,\(^{14,18}\) which very rapidly passes INT2\(^{+}\). The frequency changes during the internal \(\text{C}_6==\text{C}_5\) rotation are small. Based on the pioneering work of Langenbacher et al.,\(^{17}\) Liu et al. suggested an alternative sequential mechanism whereby the \(\text{C}_5==\text{C}_4\) splitting occurs within 10 ps and the \(\text{C}_6==\text{C}_5\) splitting occurs on a time-scale of 90 ps.\(^{17}\) This slow dynamics implies that the molecular geometry fluctuates around INT2\(^{+}\) before the \(\text{C}_6==\text{C}_5\) splitting. Trajectory B approximately mimics this mechanism.

### 2.4. FSRS Signals for the Two Model Trajectories.

The FSRS experiment starts with an impulsive actinic pump pulse \(\epsilon_p\) at time 0, which triggers the dynamics. The ultrafast Raman probe \(\epsilon_s\) comes at time \(T\) in the presence of the long-duration pump \(\epsilon_s\), and the signal is given by the change in the transmitted intensity of the probe pulse:

\[
S_{\text{FSRS}}(\omega, T) = IP^{(5)}(\omega, T)\epsilon_s^*(\omega)
\]
envelope with center frequency $\omega_3$ and duration $\sigma$, $\psi_3(t) = e^{-i(\sigma-t)/2\sigma^2-i\omega_3(t-T)}$. The polarization is then given by

$$
P^{(S)}(t, T) = \left( \frac{i}{\hbar} \right)^5 \sum_{a,b,c} |e_a|^2 |e_c|^2 \frac{IV_{ab}V_{bc}^*}{(\omega_2 - \omega_{bc})^2}
$$

$$
\times \exp \left[ (\omega_2 - \gamma_{\omega a})t \right] \int_{-\infty}^{t} d\tau_3 \exp \left[ -\left( \frac{\tau_3 - T}{2\sigma^2} \right) + i\omega_3 T + i(\omega_2 - \omega_3)\tau_3 \right]
$$

$$
+ i \int_{-\infty}^{t} \omega_a(\tau) d\tau + \gamma_{\omega a} F_3
$$

(3)

where $\omega_{ij}$ is a vibrational transition frequency in the electronic state prepared by the actinic pulse triggering electron transfer, whereas $b$ is a higher-lying excited state involved in the Raman process (Scheme 2). $V_b$ and $\omega_b$ are the transition dipole and frequency for the $i \to j$ transition. The vibrational dephasing time $\gamma_{\omega a}$ is 532 fs (line width, 10.0 cm$^{-1}$). The models for the time-dependent frequency $\omega_a(t)$ of a specific marker band are discussed in section 2.3. Other parameters are listed in Table 2.

**Table 2. Parameters Employed in the FSRS Calculations**

| $\omega_3$ (cm$^{-1}$) | $\sigma$ (fs) | $\tau_0$ (fs) | $\gamma_{\omega a}$ (fs) |
|------------------------|--------------|--------------|------------------------|
| $\omega_2 - 1800.0$    | 20.0         | 500.0        | 332.921                |
| 100000.0               |              |              |                        |

To evaluate $P^{(S)}(t, T)$, we assumed equal Raman activities for the six normal modes (i.e., set the $|e_a|^2|e_c|^2(V_{ab}V_{bc}^*)/(\omega_2 - \omega_{bc})^2$ prefactor in eq 3 to 1) and impulsive electron transfer from photoexcited FADH$^+$ to the neutral thymine dimer. In this way, we separate the pure molecular response from the superimposed initiation process in order to compute the subsequent vibrational dynamics and the microscopic splitting mechanism of the TT$^{+\gamma}$. The detailed electron transfer dynamics and its effect on the dimer signals are beyond the scope of this study. Earlier studies$^{*}$ established that the CPD-PL enzyme primarily stabilizes the FADH$^+$ excited state and slows down the futile back electron transfer.

The simulated FSRS signals for trajectory A are shown in Figure 5a. The peak positions do not resemble the instantaneous frequencies, and even the number of the peaks is higher than expected from a snapshot picture, in particular at early delay times. This reflects an interference of dispersive and absorptive line shapes$^{55}$ induced by the fast time scale $\tau_0$ of frequency changes ($\lesssim \gamma_{\omega a}^{-1}$). The time-dependent vibrational frequencies cause broadening of the peaks. Dispersive features give a clear characteristic signature of the ultrafast concerted mechanism. No clear signature of the $C_2-C_4$ splitting appears during the geometry change in region I (up to $\sim 500$ fs). When the geometry reaches region II, a dispersive line shape is found around 1560 cm$^{-1}$, which indicates the end of the spontaneous $C_2-C_4$ splitting. It originates from the fast frequency shift (brown line in Figure 3b) due to the end of both intradimer bond splittings and formation of one $C_6$ bond. As the geometry passes region III (>1000 fs), one can see the instantaneous frequencies with high spectral resolution.

The FSRS signals for trajectory B shown in Figure 5b closely resemble the instantaneous frequencies. This is because $\tau_0$ is sufficiently slow, compared to the vibrational dephasing time $\gamma_{\omega a}^{-1}$. The FSRS signal thus directly monitors the bond rearrangement of the sequential mechanism, discussed in section 2.2, at all delay times. In Figure 5b, the $C_2-C_4'$ splitting (as well as the internal $C_5-C_6$ rotation) occurs at about 100 ps and the $C_5-C_6$ splitting is reached within another 100 ps. In addition, the partial $C_2-C_4'$ splitting is completed within 10 ps, which can be observed as the blue shift of the low-frequency 1615 cm$^{-1}$ $C_3=O$ mode. The FSRS signals are sensitive to the forward electron transfer since four $C=O$ stretch frequencies are significantly red-shifted after the electron uptake (see Table S2 and Figure S12).

The snapshot limit has been observed in various systems by different spectroscopic techniques, for example, retinal isomer-
ization studied with FSRS (at sufficiently long delay times \( T > 1 \) ps),\textsuperscript{22,29} dynamics of green fluorescent protein with two-photon fluorescence spectroscopy,\textsuperscript{56} dynamics in proteins and enzymes with two-dimensional infrared (2DIR) spectroscopy,\textsuperscript{57} photo-switching of a peptide\textsuperscript{58} and biomimetic molecule\textsuperscript{59} with IR transient absorption spectroscopy. The various techniques provide different controls over the resolution, which results in different signals. The snapshot limit may be achieved with one technique but not with another. Furthermore, some of these techniques can explore larger space of evolving degrees of freedom. In particular, if two degrees of freedom (two nuclear coordinates) become coupled, fast contributions that go beyond the snapshot limit are expected. In this case, the nonadiabatic vibrational dynamics which is normally obtained in a snapshot limit for an individual coordinate becomes untraceable.\textsuperscript{60} In the following, we investigate the CPD repair mechanism and discuss the conditions whereby it can be described by a snapshot limit.

### 3. CPD SPLITTING IN THE THYMIDYLYL(3′→5′)THYMIDINE RADICAL ANION

We next turn to the entire TpT\textsuperscript{**} dinucleotide shown in Figure 1. The TpT\textsuperscript{**}, whose net charge is \(-2e\) and spin multiplicity is \(2\), is more realistic than the TT\textsuperscript{**}, and their comparison reveals the backbone effect of the DNA strand.\textsuperscript{61,62} Geometry optimization, IRC calculations, and frequency analyses along the IRC path were performed in a similar manner to TT\textsuperscript{**}. We introduce a more realistic kinetic model which takes into account the potential energy profile, and investigate the FSRS signals using the stochastic Liouville equation. We compare the resulting signals to the static averaged FSRS signals of transient intermediates.

#### 3.1. Optimized Geometries and the IRC Path: Comparison of TpT\textsuperscript{**} and TT\textsuperscript{**}

We obtained the geometries of the closed form, two INTs, three TSs, and the open form, as in the TT\textsuperscript{**}. The calculated barrier heights are in reasonable agreement with previous theoretical studies (Table 1).\textsuperscript{11,14,18} The potential energy of the TpT\textsuperscript{**} along the IRC path is shown in Figure 6. Compared with TT\textsuperscript{**}, the energy difference between INT1\textsuperscript{**} and INT2\textsuperscript{**} is smaller due to the internal \( C_5\)–\( C_6 \) rotation and a backbone distortion in INT2\textsuperscript{**}. In addition, the open form is stabilized so that both intradimer splitting processes are exothermic, which agrees with the experimental findings.\textsuperscript{67,51} The geometric changes along the IRC path and the frequency changes of marker bands of the cyclobutane fragment, shown in Figure S13, are similar to those of the TT\textsuperscript{**}. For example, the geometric change in region I corresponds to the \( C_5\)–\( C_6 \) splitting and the internal \( C_5\)–\( C_6 \) rotation, and the change in region II corresponds to the \( C_5\)–\( C_6 \) splitting. This is because in the TpT\textsuperscript{**}, the SOMO is also localized in the thymine dimer moiety and its change along the IRC path closely resembles TT\textsuperscript{**}, as seen in Figure S14.

#### 3.2. Sequential First-Order Kinetics and the FSRS Signals

The FSRS signals of TpT\textsuperscript{**} simulated with the trajectories A and B are similar to those of the TT\textsuperscript{**} (see Figure S15), except that the dispersive line shape which indicates the passage through region II is found at a higher frequency (\( \sim 1720 \) cm\(^{-1}\)) in trajectory A. We now introduce a more realistic kinetic model for the FSRS signals, which takes into account the actual potential energy profile.

We considered a sequential model based on multistep first-order kinetics and transition state theory.\textsuperscript{63} We assumed a linear sequence of intradimer bond splitting of the TpT\textsuperscript{**} with back reactions:

\[
\text{closed form} \overset{k_1}{\underset{k_{-1}}{\leftrightarrow}} \text{INT1}^{**} \overset{k_2}{\underset{k_{-2}}{\leftrightarrow}} \text{INT2}^{**} \overset{k_3}{\underset{k_{-3}}{\leftrightarrow}} \text{open form}
\]  

These satisfy the rate equation,

\[
\frac{d}{dt} \rho_{aa}(t) = -K \rho_{aa}(t)
\]

where the population \( \rho_{aa}(t) \) consists of the concentrations of the four species (e.g., \( \rho_{a(\text{closed})}(t) \)) after electron transfer due to the actinic pulse \( \varepsilon_1 \), and the state \( a \) represents the vibrational ground state. The rate matrix \( K \) was calculated with the transition state theory at room temperature (see Table S3). The solution of eq 5 is given by

\[
\rho_{aa}(t) = U \exp[-K \Delta t] U^{-1} \rho_{aa}(0)
\]

where \( U \) is a transformation matrix, which diagonalizes \( K \) into \( K^{\text{diag}} \). \( \rho_{aa}(0) \) represents the closed form. FSRS signals were calculated using the stochastic Liouville equation.\textsuperscript{64,65} This is a convenient approach for computing spectral line shapes of a quantum system coupled to a classical bath. It assumes that the system is affected by the bath, but the bath undergoes an independent stochastic dynamics that is not affected by the system. System/bath entanglement is neglected. The stochastic Liouville equations describe joint dynamics of the system and bath density matrix \( \rho \):

\[
\frac{d\rho}{dt} = \hat{L}\rho(t) = -\frac{i}{\hbar} [H, \rho(t)] + \hat{L}\rho(t)
\]

Here \( H \) is the system Hamiltonian that depends parametrically on the bath and \( \hat{L} \) is a Markovian master equation for the bath. We introduce a simple stochastic model, where the bath has four states, namely closed, INT1\textsuperscript{**}, INT2\textsuperscript{**}, and open forms in eq 4.\textsuperscript{66} The system has two vibrational states \( a \) and \( c \) and the vibrational frequency \( \omega_{ac} \) is perturbed by the bath state \( s \).
total density matrix ρ has thus 16 components |u,v⟩⟨v,u|) which represent the direct product of four Liouville space states |u,v⟩, where u,v = a,c, and four bath states s. The Liouville operator $\hat{L}$ is diagonal in the vibrational Liouville space and is thus given by four 4x4 diagonal blocks in bath space:

$$[\hat{L}]_{\nu',\nu,s,a,c} = \delta_{\nu,\nu'} \delta_{s,s'} (\nu_0 \Delta \nu', \nu_0 \Delta 
u) + \delta_{\nu,\nu'} (\nu_0 \Delta \nu', \nu_0 \Delta \nu)$$ (8)

where $\hat{L}_s = -iK$ describes the kinetics given by eq $S$ and the coherent part $\hat{L}_s = -(i/\hbar) [H_{ss'-ss},]$ describes the vibrational dynamics (see eq $S$ of Supporting Information). Following the approach outlined in ref 65, we obtain eq $S$ for the Stokes contribution to the FSRS signal.

The calculated rate for the $C_5-C_6$ splitting is slow (14.1 ps), compared to the rates of the two preceding processes (95.4 fs and 1.39 ps), and the kinetics is also closely related with the sequential mechanism proposed from experiments.6 This is clear from the time-dependent populations of the closed, INT1++, INT2++, and open forms, shown in Figure 7a. The

$$\rho_{aa}(\omega + \Delta) \approx \epsilon_1(\omega + \Delta) e^{i\Delta T \rho_{aa,aa}(-\Delta)}$$ (9)

where $\rho_{aa}(-\Delta)$ is a Fourier transform of the population of the state a given by eq $S$, $G_{aa,aa}(\omega)$ is a frequency-domain Green’s function given by eq 3, and $\sum_a$ represents the sum over species. It follows from eq 9 that the integral over $\Delta$ represents a path integral over the bandwidth corresponding to the inverse dephasing time scale (see ref 30). This integral is generally a complex quantity. Therefore, the signal given by eq 9 is governed by both real and imaginary parts of the coherence Green’s function $G_{aa,aa}(\omega)$ and thus contains dispersive features in the spectra. In the limit of slow fluctuations, one can neglect the jump dynamics during the dephasing time. In this case, replacing $\epsilon_1(\omega + \Delta) \approx \epsilon_1(\omega)$, the integral over $\Delta$ yields simply $\rho_{aa}(T)$, and we obtain the static averaged signal,

$$S_{FSRS}(\omega, T) = \sum_{a} \sum_{s} S_{FSRS,s}(\omega) \rho_{aa,aa}^{(s)}(T)$$ (10)

where

$$S_{FSRS,s}(\omega) = -\mathcal{R} \frac{2}{\hbar^2} |\epsilon_1(\omega)|^2 |\epsilon_2|^2 \sum_{c} \alpha_c^2 G_{aa,ac}(\omega - \omega_c)$$ (11)

corresponds to a linear transmission of the Raman pulse. To compute the signal given by eq 10, we first calculate individual signals of the four species, $S_{FSRS,s}(\omega)$ at each excited state. We then average over the signals with their transient concentrations $\rho_{aa,aa}(T)$ given in eq 6. The static averaged FSRS signal (eq 10) contains purely absorptive features due to the neglect of bath stochastic dynamics during the dephasing time. Furthermore, the time evolution is governed in this case by a snapshot of the populations of the excited states. The two signals given by eqs 9 and 10 are thus expected to differ at short time and be similar at longer time. Indeed after $T = 100$ fs, the stochastic Liouville equation and the static average signals (Figures 8 and 16) are virtually identical.

It is interesting how the snapshot limit is connected with nonadiabatic dynamics especially in the context of the CPD repair mechanism. Despite the fact that the mechanism of intradimer C—C bond splitting consists of several adiabatic structural changes, the peak vibrational frequency does follow instantaneously the reaction dynamics. Recently it has been shown that vibrational adiabaticity does not correspond to the
vibrations are marker bands, and their changes were rationalized by the SOMO evolution of the transient intermediates. The characteristics of the concerted and the sequential mechanism were studied by model trajectories. The difference in lifetime around the INT2− geometry, (T−T)−, and consequent differences in the FSRS signals are key signatures for distinguishing between the two proposed mechanisms. We thus demonstrated that FSRS can be a useful tool for probing the underlying molecular mechanism of the intradimer bond splitting by focusing on the time evolution of the marker bands. The direct molecular probe of the actual splitting dynamics is required in order to understand the high repair quantum yield of the photoreactivation, the delicate balance between the splitting processes and the back electron transfer, and the effect of CPD-PL protein structure on the splitting rates.

**ASSOCIATED CONTENT**

Supporting Information

Discussion on the synchronous concerted mechanism of the TT−, the changes of several molecular properties along the IRC paths of the ‘T−’ and ‘T−T’− (potential energies, key geometric parameters, Raman activities, vibrational frequencies, Duschinsky rotation matrices, and SOMOs), FSRS signals of the ‘T−T’−, the stochastic Liouville equation, and repair quantum yield calculation in sequential photoreactivation kinetics. This material is available free of charge via the Internet at http://pubs.acs.org.

**AUTHOR INFORMATION**

**Corresponding Author**

smukamel@uci.edu

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

The support of the Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy is gratefully acknowledged. We also gratefully acknowledge support from the National Science Foundation (Grant No. CHE-1361516 and No. CHE-0840513) and the National Institutes of Health (Grant No. GM-59230). B.P.F. thanks the German Research Foundation (Grant No. 2/1136/2-1) for funding within the Emmy-Noether Program.

**REFERENCES**

1. Sancar, A. Biochemistry 1994, 33, 2–9.
2. Sancar, A. Chem. Rev. 2003, 103, 2203–2237.
3. Weber, S. BBA-Bioenergetics 2005, 1707, 1–23.
4. Yang, W. Protein Sci. 2011, 20, 1781–1789.
5. Sinha, R. P.; Hader, D. P. Photochem. Photobiol. Sci. 2002, 1, 225–236.
6. Liu, Z.; Tan, C.; Guo, X.; Kao, Y.-T.; Li, J.; Wang, L.; Sancar, A.; Zhong, D. Proc. Natl. Acad. Sci. U.S.A. 2011, 108, 14831–14836.
7. Thiagarajan, V.; Byrdin, M.; Eker, A. P. M.; Muller, P.; Brettel, K. Proc. Natl. Acad. Sci. U.S.A. 2011, 108, 9402–9407.
8. Sancar, A. J. Biol. Chem. 2008, 283, 32153–32157.
9. Brettel, K.; Byrdin, M. Curr. Opin. Struct. Biol. 2010, 20, 693–701.
10. Park, H.-W.; Kim, S.-T.; Sancar, A.; Deisenhofer, J. Science 1995, 268, 1866–1872.
11. Chatgilialoglu, C.; Guerra, M.; Kaloudis, P.; Houéé-Lévê, C.; Marignier, J.-L.; Swaminathan, V. N.; Carell, T. Chem.—Eur. J. 2007, 13, 8979–8984.

![Figure 8. Variation of simulated FSRS signals of the TpT**− with different delay times T.](image-url)
(12) Hassanali, A. A.; Zhong, D.; Singer, S. J. J. Phys. Chem. B 2011, 115, 3860–3871.
(13) Hassanali, A. A.; Zhong, D.; Singer, S. J. J. Phys. Chem. B 2011, 115, 3848–3859.
(14) Masson, F.; Laino, T.; Rothlisberger, U.; Hutter, J. ChemPhysChem 2009, 10, 400–410.
(15) Kao, Y.-T.; Saxena, C.; Wang, L.; Sancar, A.; Zhong, D. Proc. Natl. Acad. Sci. U.S.A. 2005, 102, 16128–16132.
(16) MacFarlane, A. W., IV; Stanley, R. J. Biochemistry 2003, 42, 8558–8564.
(17) Langenbacher, T.; Zhao, X.; Bieser, G.; Heelis, P. F.; Sancar, A.; Michel-Beyerle, M. E. J. Am. Chem. Soc. 1997, 119, 10532–10536.
(18) Masson, F.; Laino, T.; Tavernelli, I.; Rothlisberger, U.; Hutter, J. J. Am. Chem. Soc. 2008, 130, 3434–3450.
(19) Mohammed, O. F.; Pines, D.; Dreyer, J.; Pines, E.; Nibbering, E. J. Science 2005, 310, 83–86.
(20) Schreier, W. J.; Schrader, T. E.; Koller, F. O.; Gilch, P.; Crespo-Hernández, C. E.; Swaminathan, V. N.; Carell, T.; Zinth, W.; Kohler, B. Science 2007, 315, 625–629.
(21) Kukura, P.; McCamant, D. W.; Mathies, R. A. Annu. Rev. Phys. Chem. 2007, 58, 461–488.
(22) Kukura, P.; McCamant, D. W.; Yoon, S.; Wandschneider, D. B.; Mathies, R. A. Science 2005, 310, 1006–1009.
(23) Takeuchi, S.; Ruhman, S.; Tsuneda, T.; Chiba, M.; Taketsugu, T.; Tahara, T. Science 2008, 322, 1073–1077.
(24) Yoshizawa, M.; Kurosawa, M. Phys. Rev. A 1999, 61, 013808.
(25) Fang, C.; Frontiero, R. R.; Tran, R.; Mathies, R. A. Nature 2009, 462, 200–204.
(26) Sun, Z.; Lu, J.; Zhang, D. H.; Lee, S.-Y. J. Chem. Phys. 2008, 128, 144114.
(27) Zhao, B.; Sun, Z.; Lee, S.-Y. J. Chem. Phys. 2011, 134, 024307.
(28) Lee, S.-Y.; Zhang, D.; McCamant, D. W.; Kukura, P.; Mathies, R. A. J. Chem. Phys. 2004, 121, 3632–3642.
(29) Mukamel, S.; Biggs, J. J. J. Chem. Phys. 2011, 134, 161101.
(30) Dorfman, K. E.; Fingerhut, B. P.; Mukamel, S. Phys. Chem. Chem. Phys. 2013, 15, 12348–12359.
(31) Umapathy, S.; Lakshmanan, A.; Mallick, B. J. Raman Spectrosc. 2009, 40, 235–237.
(32) Fingerhut, B. P.; Dorfman, K. E.; Mukamel, S. J. Phys. Theory Comput. 2014, 10, 1172–1188.
(33) Fingerhut, B. P.; Dorfman, K. E.; Mukamel, S. J. Chem. Theory Comput. 2013, 4, 1933–1942.
(34) In the TT™, unlike in the TpT™, there is no unique open form due to the lack of DNA backbone. The end-point geometry of an IRC calculation is thus regarded as the open form.
(35) Fukui, K. Acc. Chem. Res. 1981, 14, 363–368.
(36) A displacement between two molecular structures was evaluated by positioning them in the same xyz frame so as not to generate any linear or angular momentum. This method was, for example, employed to evaluate a distance between an optimized geometry and an end-point geometry of IRC calculation, which is useful to connect several IRCs.
(37) Seeger, R.; Pople, J. A. J. Chem. Phys. 1977, 66, 3045–3050.
(38) Buevanschmitt, R.; Ahlrichs, R. J. Chem. Phys. 1996, 104, 9047–9052.
(39) Hohre, W. J.; Ditchfield, R.; Pople, J. A. J. Chem. Phys. 1972, 56, 2257–2261.
(40) Hariharan, P. C.; Pople, J. A. Theor. Chim. Acta 1973, 28, 213–222.
(41) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. J. Chem. Theory Comput. 2006, 2, 364–382.
(42) Zhao, Y.; Truhlar, D. G. Acc. Chem. Res. 2008, 41, 157–167.
(43) Fingerhut, B. P.; Heil, K.; Kaya, E.; Oesterling, S.; de Vivie-Riedle, R.; Carell, T. Chem. Sci. 2012, 3, 1794–1797.
(44) Hohenstein, E. G.; Chill, S. T.; Sherrill, C. D. J. Chem. Theory Comput. 2008, 4, 1996–2000.
(45) Jurčeka, P.; Sponer, J.; Černý, J.; Hobza, P. Phys. Chem. Chem. Phys. 2006, 8, 1985–1993.