Indications of 5′ to 3′ Interbase Electron Transfer as the First Step of Pyrimidine Dimer Formation Probed by a Dinucleotide Analog

Yajun Jian,[†] [a, b] *Egle Maximowitsch,*[†] [c] *Surya Adhikari,*[a] *Lei Li,**[a,d] *and Tatiana Domratcheva**[c]

Abstract: Pyrimidine dimers are the most common DNA lesions generated under UV radiation. To reveal the molecular mechanisms behind their formation, it is of significance to reveal the roles of each pyrimidine residue. We thus replaced the 5'-pyrimidine residue with a photochemically inert xylene moiety (X). The electron-rich X can be readily oxidized but not reduced defining the direction of interbase electron transfer (ET). Irradiation of the XpT dinucleotide under 254 nm UV light generates two major photoproducts: a pyrimidine (6-4) pyrimidone (6-4PP) analog and an analog of the so-called spore photoproduct (SP). Both products are formed via reaction at C4=O of the photo-excited 3'-T, which indicates that excitation of a single “driver” residue is sufficient to trigger pyrimidine dimerization. Our quantum-chemical calculations demonstrated that photo-excited 3'-T accepts an electron from 5′-X. The resulting charge-separated radical pair lowers its energy upon formation of interbase covalent bonds eventually yielding 6-4PP and SP.

Introduction

As the genetic information carrier, DNA is under constant environmental stresses, such as UV radiation from the solar light. Among the four 2′-deoxyribonucleotides, thymidine is the most UV sensitive one followed by 2′-deoxycytidine.[1] UV radiation of pyrimidine residues leads to dimerization reactions, with cyclobutane pyrimidine dimers (CPDs) and pyrimidine (6-4) pyrimidone photoproducts (6-4PPs) as the two most common photoproducts (Figure 1).[1a, 1c-e] The pyrimidone moiety in 6-4PP absorbs at 326 nm,[2] that can be excited by UVA/UVB light, leading to the formation of Dewar valence photoproduct (Dewar PP) via a ring rearrangement reaction.[3] Moreover, in bacterial endospores where the genomic DNA is suggested to adopt an A-like conformation, UV radiation results in another thymidine dimer, 5-thyminyl-5,6-dihydrothymine, which is commonly referred to as the spore photoproduct (SP).[4] It is generally believed that SP is mainly formed in spores, whereas CPD, 6-4PP and Dewar PP are formed in a wide range of species. In humans, mutations induced by pyrimidine dimers are considered as the dominant cause for the occurrence of skin cancer.[5] It is therefore of significance to understand the pyrimidine photochemistry in detail.

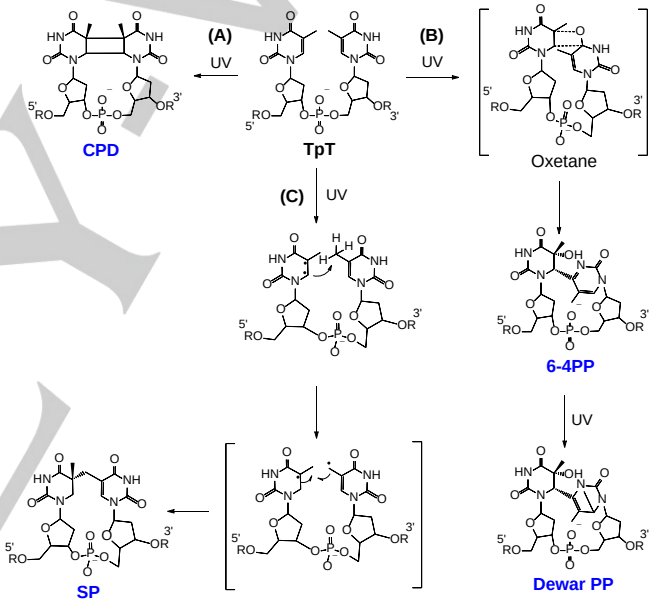


Figure 1. Thymidine dimerization reactions under UV radiation. (A) Formation of CPD. (B) Formation of 6-4PP via a putative oxetane intermediate. The formed 6-4PP can further isomerize into Dewar PP under UVA/UVB light. (C) Formation of SP which is initiated by an H-atom abstraction mediated by a diradical species formed at the C5=C6 bond followed by radical recombination.

Despite the fact that CPDs, 6-4PPs and SP have been discovered for half a century or so,[6] issues still remain regarding their formation mechanisms. The formation of CPDs is indicated to occur via a [2 + 2] cyclophotoaddition reaction. Both pyrimidine singlet and triplet excited states are found to be able to induce CPD formation.[7] Between these two states, the triplet states represent a minor reaction channel and only contribute to < 10% of CPDs formed in single-stranded DNA under UVC (100 – 290 nm) and UVB (290 – 320 nm) irradiation.[8] Surprisingly, the long-wavelength UVA light (320 – 400 nm) was also found to induce CPD formation potentially through triplet energy transfer.[1e, 9] CPDs were found to be generated after finishing exposure to UVA light presumably by a chemiexcitation

[a] Dr. Yajun Jian, Dr. Surya Adhikari, Prof. Lei Li
Department of Chemistry and Chemical Biology
Indiana University-Purdue University Indianapolis (IUPUI)
402 North Blackford Street, Indianapolis, Indiana, 46202, United States
E-mail: lilei@iupui.edu
[b] Dr. Yajun Jian
School of Chemistry & Chemical Engineering
Shaanxi Normal University (SNNU)
No. 620, West Chang'an Avenue, Xi'an, Shaanxi, 710119, P. R. China
[c] Egle Maximowitsch, Dr. Tatiana Domratcheva
Department of Biomolecular Mechanisms
Max-Planck Institute for Medical Research
Jahnstrasse 29, 69120 Heidelberg, Germany
E-mail: Tatjana.Domratcheva@mpimf-heidelberg.mpg.de
[d] Prof. Lei Li
Department of Dermatology
Indiana University School of Medicine
Indianapolis, Indiana 46202, United States

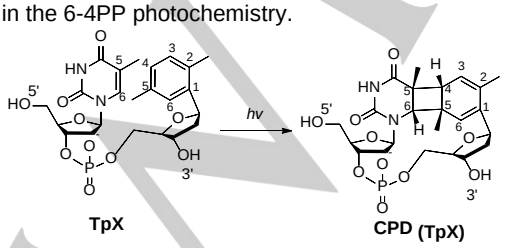
[†] The authors contributed equally to this work.

Supporting information for this article is given via a link at the end of the document.

mechanism.[10] 6-4PP is indicated to occur via an oxetane or azetidine intermediate formed between the two involving pyrimidine residues after their photo-excitation by UVC/UVB light; such an intermediate then thermally decomposes into the corresponding 6-4PP.[1a] UVA light is unable to induce 6-4PP formation, indicating that 6-4PP is formed via pyrimidine singlet excited states.[1d, 9a] SP formation is achieved via a diradical mediated H-atom abstraction followed by radical recombination mechanism.[11] In the presence of pyridopsoralens as photosensitizer, UVA light can also induce SP formation,[12] suggesting that thymine triplet excited state is likely responsible for the SP photochemistry.[13]

A recent computational study with dinucleotide TpT implies that the excitation of the 3'-T alone may be sufficient for the 6-4PP photochemistry.[14] The 3'-T is excited to the triplet state displaying an elongated C4=O bond via a fast singlet−triplet crossing, which then reacts with the C5=C6 bond of the 5'-T to form 6-4PP.[14] This result contradicts with the previous finding that only the singlet excited state is involved in the 6-4PP formation;[1d] the reason for this discrepancy is unclear. Further computational studies suggest the 6-4PP photochemistry is triggered by the so-called charge-transfer states featuring electron transfer (ET) from the 5'-base to the 3'-base in the TpT and TpC sequences.[7, 15] However, given that the two pyrimidine residues likely possess similar redox potentials, ET in both directions, from 5′ to 3′ and from 3′ to 5′, should be possible. Therefore, there is an urgent need to identify an experimental system by replacing a pyrimidine with a photo-inert but ET-active moiety to test these computational results above.

Because pyrimidine photoreactions are largely governed by the base stacking conformation,[16] the replacement should possess a similar shape and stacking interaction to a natural pyrimidine. Toluene and its derivatives like 2,4-difluorotoluene are close steric analogues of thymine;[17] they can be incorporated into DNA by polymerases as if they were thymines.[17c, 18] The incorporations do not lead to drastic DNA conformational changes, as proved by NMR spectroscopic[19] and crystallographic studies.[20] However, these residues possess rather different electronic structures from that of thymine due to their higher symmetries, absence of the $n\pi^*$ states and lack of tautomeric structures,[1a, 21] making them rather photo-inert. At the same time, their aromatic rings are electron-rich, that facilitates electron donation but impedes electron acceptance, thus determining the direction for any possible interbase ET. These properties render toluene or its derivatives as a perfect thymine substitute to reveal mechanistic details in the 6-4PP photochemistry.



Scheme 1. TpX photoreaction yielding a CPD analog.

Previously, we replaced the 3'-T with a toluene (To) or a *meta*-xylene (X) group, the resulting dinucleotide TpTo or TpX supports CPD formation under UVC light (Scheme 1).[22] In this report, we retained the 3'-T and replaced the 5'-T by an X to obtain an XpT dinucleotide to shed light on the 6-4PP formation because under UV light, the xylene oxidation is preferred in XpT resulting a pair of 5′-X cation and 3′-T anion radicals. The XpT photoreaction indeed affords a 6-4PP analog as a major product, providing direct experimental evidence that the photo-excited 3′-residue alone is enough to trigger the 6-4PP formation and the reaction proceeds via a 5′ to 3′ ET process. The reaction also yields a SP analog; while no CPD analogs are obtained. Our quantum-chemical calculations further demonstrate that the interbase ET from the 5′-X to the 3′-T is responsible for the formation of both photoproducts; the 6-4PP and SP formation may thus share some similar mechanistic scheme.

Results

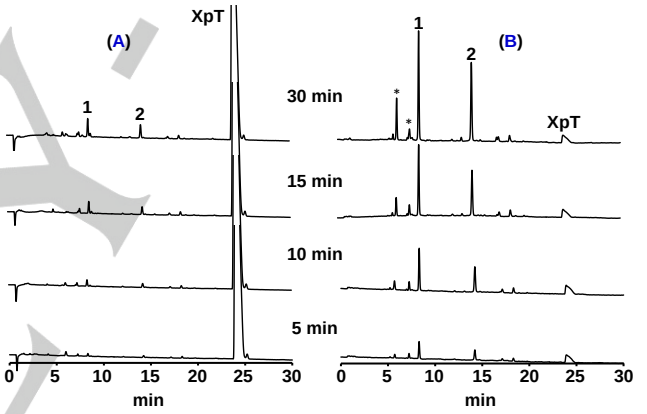


Figure 2. HPLC chromatograms monitored at 230 nm (A) and 320 nm (B) for the XpT photoreaction initiated under 254 nm UV light in 10 mM pH 7.4 phosphate buffer at ambient temperature. XpT was irradiated for 5, 10, 15 and 30 minutes; two thermally stable products **1** and **2** were generated as the major species, both of which absorb at 320 nm. The XpT photoreaction also generated several minor species. Using mass spectrometry, the compounds donated with * appear to be water adducts, the structures of which were not pursued in the current study.

Product formation in XpT photoreaction. UV radiation of the XpT solution resulted in two major products **1** and **2** that are thermally stable at ambient temperature (Figure 2). The quantum yields of **1** and **2** were determined as $(3.2 \pm 1.0) \times 10^{-4}$ and $(3.3 \pm 1.0) \times 10^{-4}$ respectively. ESI-MS analyses at the negative ion mode revealed that **1** possesses an *m/z* value of 525.17 for the $[M - H]^-$ species, which is the same to that of XpT.[23] Formation of these species is consistent with the general property of pyrimidine photoreactions where pyrimidine residues dimerize into species with identical molecular weight to the unreacted dinucleotides. **2** exhibits a $[M - H]^-$ signal of 507.16,[23] which is 18 amu less than that of XpT, suggesting an elimination of a water molecule during its formation. Several minor products were also generated during the course of the reaction, which are likely XpT water adducts as indicated by

mass spectrometry (+ 18 amu relative to XpT in regular water and + 20 amu in ^{18}O labeled water).

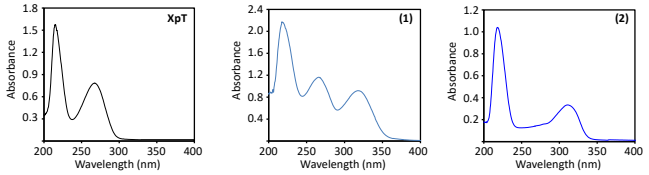


Figure 3. UV-vis spectra of XpT, **1** and **2** in aqueous solution at pH 7.0. **1** and **2** exhibit absorption bands at 320 and 317 nm respectively, indicating the existence of a pyrimidone ring in these two products.

Characterization of 1 and 2 via UV-vis spectroscopy. Also as shown in Figure 2, when monitored at 320 nm on HPLC, both **1** and **2** exhibit significant absorption while XpT barely absorbs. We obtained the UV-vis spectra for XpT, **1** and **2** (Figure 3) to shed light on their possible chemical structures. XpT does not exhibit any absorption peak at > 300 nm, while both **1** and **2** absorb around 320 nm. Because an absorption band around 320 nm is typically induced by the pyrimidone moiety formed in the 6-4PP species,[2] the UV-vis spectra indicate that such a moiety may be formed in both products.

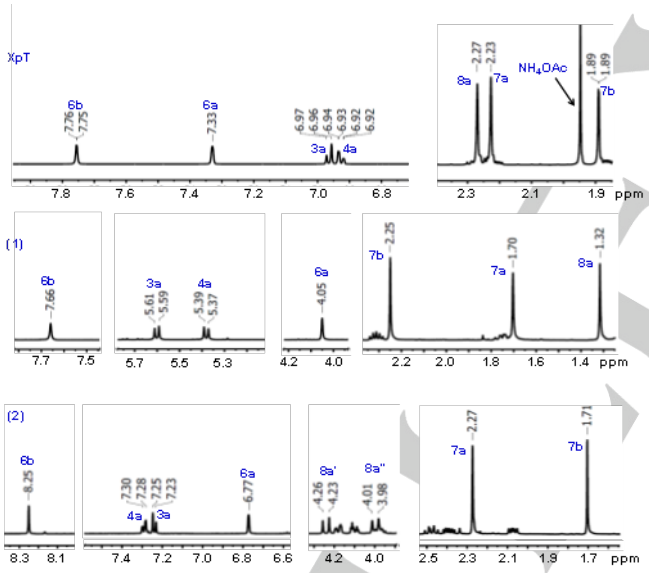


Figure 4. Zoom-in-view of the ^1H NMR spectra of XpT, **1** and **2** showing signals for protons associated with non-deoxyribose carbons. D_2O was used as the NMR solvent for **2**; MeOH-d_4 was used for the other two compounds.

Characterization of 1 via NMR spectroscopy. We next turned to NMR spectroscopy for structural analysis. The ^1H NMR spectrum of **1** shows all three methyl groups at 7a, 7b, and 8a are retained (Figure 4). For hydrogen atoms on the 5′-xylene ring, the two doublet signals at 5.60 and 5.38 ppm exhibit the same $J_{H\text{-}H}$ coupling constant; they were assigned as H3a and H4a respectively. The relatively downfield chemical shifts imply that they are still attached to unsaturated carbons. A singlet signal of 4.05 ppm was found for H6a. Comparing with signals found in XpT, these ^1H NMR signals revealed in **1** move toward the upfield region, indicating a loss of aromaticity at the xylene ring. The biggest chemical shift change was found for H6a (> 3ppm), which, together with the chemical shift change for the methyl group 8a, indicates that the reaction likely occurs at the C5a-C6a bond. Thus, similar to the CPD formation in TpX photoreactions,[22] the xylene ring in XpT also undergoes a photo-addition reaction.

Table 1. Summary of the ^1H NMR spectra for non-deoxyribose protons in XpT, **1** and **2**.

Chemical Shift (ppm)	XpT[a]	1	2
H3a	6.96	5.60	7.24
H4a	6.92	5.38	7.29
H6a	7.33	4.05	6.77
H6b	7.75	7.66	8.25
-CH_3 (7a)	2.23	1.70	2.27
-CH_3 (8a) (–CH_2- for **2**)	2.27	1.32	4.00, 4.25
-CH_3 (7b)	1.89	2.25	1.71

a. The NMR spectra of XpT can be found in the Supporting Information.

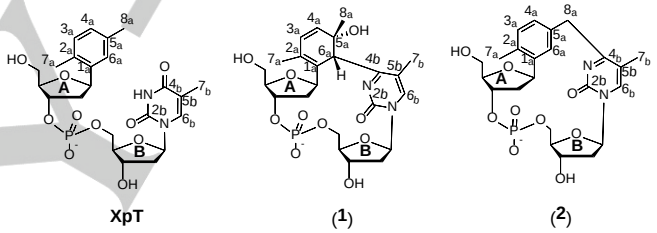


Figure 5. Chemical structures of XpT, **1** and **2** indicated by UV-vis and NMR spectroscopy. All non-deoxyribose carbons are numbered.

The methyl group 7b (2.25 ppm) and H6b (7.66 ppm) at the 3′-thymine ring are still associated with unsaturated carbons. In the HMBC spectrum, H6a and C5b exhibit a strong coupling signal (J_3), indicating a direct connection between C6a and C4b.[23-24] The chemical shifts of the four carbons on the 3'-nucleobase of **1** (also in **2** as shown below) are nearly identical to those at the pyrimidone ring in the ^{13}C NMR spectrum of the natural 6-4PP.[23, 25] Moreover, the 3′-residue does not adopt a Dewar structure, in which H6b exhibits a chemical shift of 5.28 ppm and C6b exhibits a chemical shift of 72.3 ppm due to the saturation at this hydrocarbon.[3] Therefore, we conclude that **1** possesses a 3'-pyrimidone ring and is a 6-4PP analog (Figure 5), agreeing with the UV-vis data.

In the ROESY spectrum, H6a interacts with H1a' and H4a' but not with H2a' and H3a' at the 2′-deoxyribose, implying that the 5'-nucleoside adopts a syn conformation. H6b exhibits strong interactions with H2'b and H3'b, suggesting an anti conformation for the 3'-nucleoside.[26] Similar interactions were observed in the natural 6-4PP TpT.[2, 24, 26b] However, since the xylene ring has no carbonyl moiety at C4a, the absolute configuration at C5a is *S* not *R* found in the natural 6-4PP; both C5a and C6a in **1** adopt an *S* configuration.

Characterization of 2 via NMR spectroscopy. In the ^{1}H NMR spectrum, one of the methyl moieties (8a) disappeared and new signals composed of a pair of doublets with a large coupling constant ($J_{H\text{-}H}$ = 15.5 Hz) emerged at 4.00 and 4.25 ppm (Figure 4, Table 1), indicating the formation of a methylene group at C8a. Moreover, the H8a protons exhibit strong coupling interactions (J_2) with C4b and C5a, indicating C8a is directly linked to C4b. The xylene ring in **2** remains intact, as implied by the ^{1}H NMR signals similar to those found in XpT. H6b (8.25 ppm) is still located on an aromatic ring, which is consistent with the formation of a 3′-pyrimidone moiety. Therefore, the NMR data supports the structure of **2** shown in Figure 5.

In the ROESY spectrum, H6a exhibits crosspeaks with H2a', H3a' and H5a' at the 2′-deoxyribose, suggesting that the 5'-nucleotide adopts an anti conformation.[26] In contrast, H6b interacts with H1b', H4b' and H5b', but not with H2b' and H3b′, suggesting a syn conformation for the 3'-pyrimidone. Therefore, the two nucleotides in **2** exhibit an anti-syn combination, different from the syn-anti combination found in **1**.

Table 2. Singlet-state excitation energies (kcal/mol) and oscillator strength (in brackets) in the XpT and TpT dinucleotides computed with the XMCQDPT2-PCM method at the B3LYP-D3-optimized geometries.

Dinucleotide	$^1T^*$	$^1(X^+T^-)$ or $^1(T^+T^-)$	$^1(T^-T^+)$
XpT[a]	109.9 (0.416)	121.3 (0.045)	
TpT[b]	5': 108.3 (0.331)	132.8(0.011)	136.9 (0.003)
	3': 110.5 (0.538)		

a, Results obtained with XMCQDPT2-SA5-CASSCF(6,4)-PCM ; b, results obtained with XMCQDPT2-SA6-CASSCF(4,4)-PCM. Complete data are presented in the Supporting Information Table S1.

Table 3. Electron affinities and ionization potentials (kcal/mol) of xylene (X), thymine (T) and triplet thymine (^{3}T*) estimated using the (U)B3LYP-D3-PCM method.

Base	Electron affinity	Ionization potential
X	4.6	134.3
T	30.6	137.8
^{3}T*	98.2[a]	70.2[a]

a, The energies of ^{3}T* differ from the energies of T by the adiabatic excitation energy of ^{3}T* equal to 67.6 kcal/mol.

Energies of excited states in stacked XpT and TpT. To elucidate how **1** and **2** are formed, we performed quantum-chemical calculations. We considered a stacked geometry of the XpT dinucleotide with both sugar residues adapting the C3'-endo configuration, which is rather similar to the stacked geometry of the TpT dinucleotide fragment in B-DNA. The B-DNA-like conformation has the lowest energy among several tested conformations of XpT according to our B3LYP-D3-PCM computations (Supporting Information Figure S1). The high tendency of the XpT dinucleotide to form a stacked conformation was confirmed by our MD simulations (Supporting Information Figure S2). At the optimized stacked structures of TpT and XpT, we performed calculations of the excited-state energies and oscillator strengths with the XMCQDPT2 method supplemented with the equilibrium PCM treatment of the water solvent (below referred to as XMCQDPT2-PCM). The electronic structure of the states was identified by considering transition electron densities (Supporting Information Figure S3). In the dinucleotide, the lowest-energy state with a considerable oscillator strength giving rise to the longest-wavelength absorption band corresponds to the $\pi\pi^*$ excited thymine, the $^1T^*$ state. In XpT, the $^1T^*$ state at 109.9 kcal/mol (corresponds to 260 nm) is likely to be populated in our experiments using the 254-nm excitation-energy source. At 121.3 kcal/mol in XpT and at 132.8 kcal/mol in TpT, delocalized states that correspond to electron transfer (ET) from the 5′ base to the 3′ base, the $^1(X^+T^-)$ and $^1(T^+T^-)$ states, were found. The variation of the ET energy in the two compounds is consistent with the ionization potentials of the xylene and thymine electronic systems (Table 3). In TpT, we also identified a delocalized state that corresponds to ET from 3'-T to 5'-T, the $^1(T^-T^+)$ state, with the energy 136.9 kcal/mol.

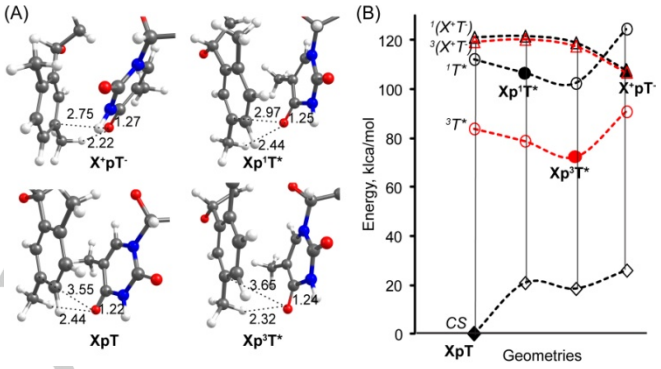


Figure 6. Interbase ET in the excited XpT dinucleotide. (A). Ground-state- and excited-state-structures optimized using the TD-B3LYP-D3 and UB3LYP-D3 methods. A dinucleotide fragment comprising the two bases is shown. Selected distances (Å) are indicated. (B). XMCQDPT2-SA9-CASSCF(6,4)-PCM energies at the optimized geometries. At each geometry, the energies of nine states (five singlets and four triplets) were computed (Supporting Information Table S2). The five lowest-energy states are included in the graph. Black and red lines connect the energies corresponding to the same electronic structure labelled in Italic font. The energy corresponding to the optimized geometry labelled in bold font is indicated with the filled symbol. For each geometry, vertical excitation energies are indicated with open symbols connected by a vertical line. The electronic structure of the states was identified from the electron density analysis (Supporting Information Figure S3).

Excited-state geometries of XpT. The computed excitation spectrum suggests population of the singlet-excited thymine Xp^{1}T*, triplet-excited thymine Xp^{3}T* and also charge-separated radical pair X$^+$pT$^-$ upon photoexcitation of XpT. Using the TD-B3LYP-D3 calculations, the optimized excited-state geometries Xp^{1}T* and X$^+$pT$^-$ were obtained; whereas using the U-B3LYP-D3 calculations, the Xp^{3}T* geometry was obtained (Figure 6A). Because of the low-energy $n\pi^*$ state of thymine,[14] the optimized Xp^{1}T* geometry corresponds to the $^1n\pi^*/^1T^*$ state crossing rather than to the true $^1T^*$-state energy minimum. The Xp^{1}T* and Xp^{3}T* geometries indicate that the $\pi\pi^*$ excited thymine adapts a non-planar geometry in the singlet and triplet electronic configurations. In contrast, the radical cation and anion X$^+$ and T$^-$ fragments of X$^+$pT$^-$ assume planar geometries. The computed net charges are +0.8 and –1.2 on X$^+$ and T$^-$, respectively. In Xp^{1}T* and Xp^{3}T*, similar to the ground-state XpT, the net charges on X and T are –0.1 and –0.3, respectively. In the T$^-$ radical anion, the C4b=O4b bond becomes elongated to 1.27 Å;

whereas the elongation of the C4b=O4b bond is less pronounced in the $\pi\pi^*$-excited thymine than in the radical anion. Furthermore, the interbase distances become shorter in X^+pT^-. In particular, the distance between C5a and O4b decreases (favorable for the formation of **1**), and the distance between the methyl group bond to C8a and O4b decreases (favorable for the formation of **2**).

Energies and electronic coupling controlling interbase ET. At the optimized excited-state geometries, we compared the energies of the singlet and triplet states computed with the XMCQDPT2-PCM method (Figure 6B). The Xp^1T^* geometry in the $^1T^*$ state has an energy of 106.6 kcal/mol. This energy is higher than the energy of the true energy-minimum because of the incomplete geometry optimization; the $^1T^*$-energy of 102.4 kcal/mol at the triplet geometry Xp^3T^* is probably closer to the energy of the true $^1T^*$-state energy minimum. The charge-separated radical pair X^+pT^- in the $^1(X^+T^-)$ and $^3(X^+T^-)$ states has energies 107.5 and 106.9 kcal/mol, respectively, which is only slightly higher than the energy of the singlet-excited thymine. The energy of Xp^3T^* geometry in the $^3T^*$ state is only 72.4 kcal/mol. This low energy indicates that formation of the charge-separated radical pair X^+pT^- is rather unfavorable in the triplet $^3T^*$ state starting from the Xp^3T^* geometry. In contrast, radical-pair formation in the singlet state starting from the Xp^1T^* geometry is energetically possible, especially upon further stabilization of the afforded radical-pair complex. The energy of the X^+pT^- optimized geometry being lower than the vertical excitation energy of thymine at the XpT ground-state geometry is also favorable for interbase ET in the singlet-excited state.

Because of the low oscillator strengths, direct population of the $^1(X^+T^-)$ state leading to the formation of the radical pair X^+pT^-. upon photoexcitation is less probable as compared to the population of the excited thymine. However, excited thymine may undergo an interbase ET reaction, accepting an electron from xylene. Indeed, electron affinity increases and the ionization potential decreases in the excited thymine (Table 3). Thus, thymine photoexcitation promotes ET in both directions in the TpT dinucleotide. In XpT, excited thymine is coupled to xylene, which is likely to serve as electron donor, therefore ET from 5'-X to 3'-T is more favorable.

In addition to the favorable energies, the rate of interbase ET depends on the amount of electronic coupling between the $^1T^*$ and $^1(X^+T^-)$ states. We estimated the electronic coupling from the excitation energies, state dipole moments and transition dipole moments according to the Mulliken-Hush method[27] (Supporting Information Table S3). The obtained coupling of the $^1T^*$ and $^1(X^+T^-)$ states at the Xp^1T^* geometry equals 69 meV. The coupling computed for the $n\pi^*$-excited thymine with the $^1(X^+T^-)$ state of 17 meV is also substantial, indicating that population of the $n\pi^*$-state thymine should not abolish interbase ET. The estimated amounts of coupling correspond to the maximum ET rate constant (assuming the activation energy is zero)[28] of 15 ps^{-1} (^{1}T* state) and 1 ps^{-1} ($n\pi^*$ state). For comparison, the electronic coupling of the $^1T^*$ and $^1(T^+T^-)$ states at the ground-state TpT geometry is 45 meV corresponding to the maximum ET rate constant of 6 ps^{-1}.

Computed absorption spectra of XpT and its photoproducts. To validate our computational procedure and to provide further support to the identified structures of **1** and **2**, we present the computed absorption spectra of these compounds (Figure 7). The geometries of **1** and **2** were optimized using the B3LYP-D3 method. Then, excitation spectra were computed with the XMCQDPT2-PCM method (Supporting Information Table S4). We note a remarkable agreement of the computed spectra (Figure 7) with the experimental spectra (Figure 3).

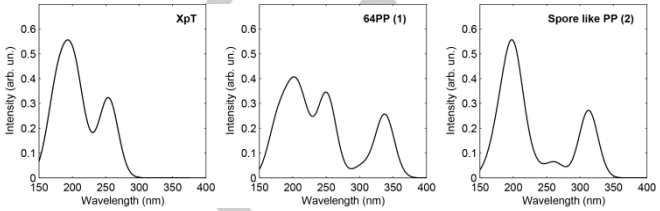


Figure 7. Computed UV-vis spectra of XpT and its major photoproducts **1** and **2**. Excited-state computations were performed with the XMCQDPT2-PCM method at the B3LYP-D3-optimized geometries. The excitation energies and oscillator strengths were convoluted with Gaussian functions of 16-nm FWHM. Complete data are listed in the Supporting Information Table S4.

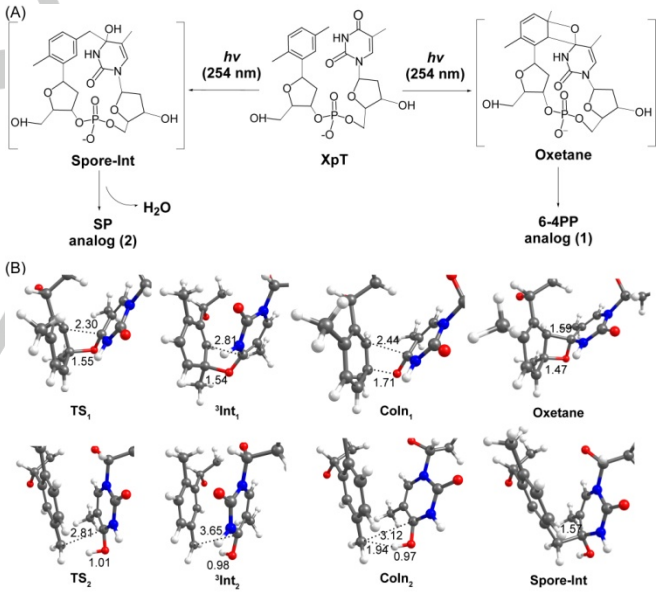


Figure 8. Formation of **1** and **2** form XpT. (A). Chemical structures of the XpT reactant and ground-state intermediates Oxetane and Spore-Int. (B) Key structures along the reaction pathways according to the (U)B3LYP-D3 calculations (the respective energies are shown in the Supporting Information Figure S4). The $CoIn_1$ and $CoIn_2$ structures were optimized with the SA2-CASSCF(12,12) method. Only a fragment of the dinucleotide comprising the two bases is shown. Selected distances (Å) are indicated.

Photochemical reaction coordinates in the ground- and triplet states. To elucidate how **1** and **2** are formed, we performed calculations of the respective reaction coordinates in the ground closed-shell (*CS*) state and in the first excited triplet state, using the (U)B3LYP-D3 method. These rather straightforward computations enabled us to identify the

structures of transition states and intermediates that are involved in the formation of interbase covalent bonds (Supporting Information Figure S4). We found that **1** and **2** are formed via the Oxetane and Spore-Int intermediates, respectively (Figure 8A). Two high-energy saddle-points TS_1 and TS_2 were found in the ground state (Figure 8B). The TS_1 geometry features an already-formed C5a−O4b bond and a 2.30-Å distance between C6a and C4b, corresponding to an activated (pre-formed) bond. Further shortening of the C6a−C4b distance leads to Oxetane. The TS_2 geometry features a formed O4b−H bond whereas the distance between C8a and C4b is found at 2.81 Å, which further decreases in Spore-Int. The Oxetane pathways are remarkably similar in XpT and TpT, even though the Oxetane energy is lower in TpT (Supporting Information Figure S5).

In the triplet-state *^3T**, formation of the interbase covalent bonds occurs in a step-wise fashion via a biradical intermediate. Formation of the C5a−O4b and O4b−H bonds leads to the biradical intermediates $^3Int_1$ and $^3Int_2$, respectively (Figure 8B). In $^3Int_1$ and $^3Int_2$, the spin density becomes delocalized over the two bases, indicating an interbase ET process. Formation of the intermediates requires more than 10 kcal/mol activation energy (Supporting Information Figure S4), which is in agreement with the previous findings for TpT.[13-14, 29] Subsequent formation of the C6a−C4b and C8a−C4b bonds results in recombination of the biradical intermediates, eventually yielding Oxetane and Spore-Int in the ground state.

Excited-state energies along the photochemical reaction coordinates. To demonstrate the key role of 5' to 3' interbase ET in the XpT photochemistry, we compared the energies of the *^1T**, *^3T**, *$^1(X^+T^-)$* and *$^3(X^+T^-)$* states along the reaction coordinates. The singlet-state route (Figure 9A) starts from the radical-pair reactant geometry X^+pT^- (107.5 kcal/mol) with the shortened C5a−O4b and H−O4b distances. Further decrease of the distances and formation of the respective covalent bonds lowers the energy of the *$^1(X^+T^-)$* state. At the same time, the energy of the *CS* ground state significantly increases, leading to conical intersections (CoIn) characterized by the degenerate *CS* and *$^1(X^+T^-)$* states. To characterize energy barriers separating the X^+pT^- minimum and the CoIn structures, we computed relaxed energy scans in the *$^1(X^+T^-)$* state by gradually decreasing the C5a−O4b and H−O4b distances and optimizing all other geometry parameters using the TD-B3LYP-D3 method (Supporting Information Figure S6). Two energy barriers, indicated as (1) and (2), were found (Figure 9A). The energies at the geometries comprising the scans were recomputed with the XMCQDPT2-PCM method. At this level of theory, only structure (2) remains an energy barrier, whereas the relative energy of structure (1) in the *$^1(X^+T^-)$* state decreases, indicating a barrierless path from the X^+pT^- minimum to the $CoIn_1$ structure.

The minimum-energy CoIn geometries (Figure 7B) were obtained using the CASSCF geometry optimization. $CoIn_1$ mediating formation of Oxetane is structurally very similar to an analogous structure previously reported for two thymines.[30] $CoIn_2$ corresponding to the Spore-Int formation is the first structure of such a type characterized in our study for the first time. $CoIn_1$ and $CoIn_2$ structurally resemble the ground-state TS_1 and TS_2 geometries. From the $CoIn_1$ and $CoIn_2$ geometries, further shortening of the C6a−C4b and C8a−C4b distances leads to the formation of Oxetane and Spore-Int. We note that the *^1T** state shows a similar energy trend as the *CS* state, which indicates that formation of Oxetane and Spore-Int in the *^1T** state requires high activation energies. Thus, thymine photoexcitation as such does not catalyze formation of the photoproducts. Instead, excited thymine undergoes interbase ET which catalyzes the formation of the photoproducts.

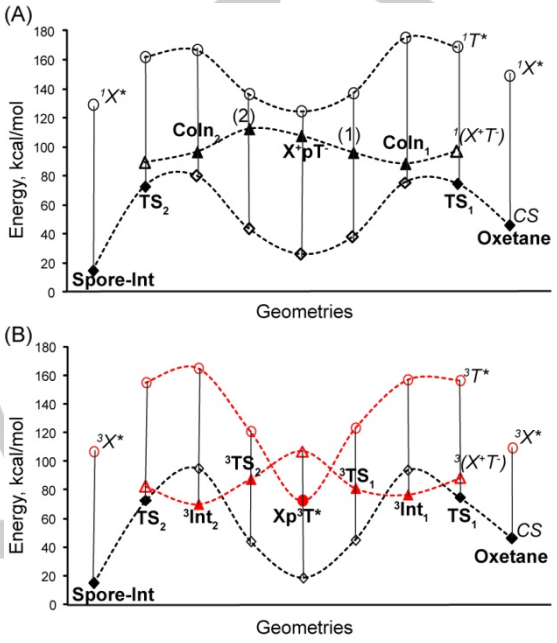


Figure 9. Excited-state energies at the key structures from the photoreaction coordinates. Only states of interest are included in the graph. (A) Energy change along the singlet-excited state route. (B) Energy change along the triplet-state route. The energy calculations and presentation style are similar to those of Figure 6B. The geometries at which the energies were computed are presented in Figure 8 and Supporting Information Figure S4. Geometries (1) and (2) correspond to the energy maxima of the relaxed energy scans computed with the TD-B3LYP-D3 method presented in the Supporting Information Figure S6. All energies are listed in the Supporting Information Table S5.

The triplet route (Figure 9B) starts from the triplet reactant Xp^3T^* (72.4 kcal/mol), which is formed from the initially populated *^1T** state via an intersystem crossing (ISC) involving a *$n\pi^*$* excited state.[14] Interbase ET in the triplet state leads to the formation of $^3Int_1$ and $^3Int_2$ (73.2 and 69.8 kcal/mol, respectively). The energy of the interbase ET state *$^3(X^+T^-)$* decreases upon formation of the C5a−O4b and O4b−H bonds, so it becomes the lowest-energy triplet state at $^3Int_1$ and $^3Int_2$. Thus, the calculations demonstrated that both singlet and triplet routes involve interbase ET from 5'-X to 3'-T. The considerable *$^3T^*$–$^3(X^+T^-)$* excitation energy and reordering of the states at the Xp^3T^*, $^3Int_1$ and $^3Int_2$ geometries clearly indicate the existence of energy barriers separating Xp^3T^* from $^3Int_1$ and $^3Int_2$. Therefore, formation of the C5a−O4b and O4b−H covalent bonds is delayed in the triplet state as compared to the singlet-excited state. At the $^3Int_1$ and $^3Int_2$ geometries, the small singlet-triplet

energy difference facilitates biradical recombination via ISC. However, recombination before the C6a−C4b and C8a−C4b bonds are formed favors decay to XpT instead of formation of Oxetane and Spore-Int, indicating that the photoproducts are unlikely to form through the triplet route.

Rearrangement of the Oxetane intermediate to 6-4PP via protonation at O4b. Despite the extensive effort to look for the Oxetane and Spore-Int intermediates, neither species was observed in our experiments. Previously, the assistance of the solvent water molecules has been suggested to accelerate Oxetane rearrangement to the final 6-4PP photoproduct via proton transfer from N3bH to O4b.[31] Here we extend this earlier hypothesis by considering the rearrangement of Oxetane via protonation of the O4b position which leads to the 6-4PP isomer in which O4bH and N3bH are simultaneously protonated. We compared protonation in the XpT Oxetane and a stable thietane using the B3LYP-D3 calculations of the respective reaction coordinates (Supporting Information Figure S7). Depending on the nature of the proton donor, opening of the Oxetane/thietane ring by O4b/S4b protonation can be energetically favorable. In particular, when the NH_4^+ molecule is considered as a proton donor, the O4b/S4b protonation is practically barrierless. The resulting protonated 6-4PP isomer is by 24 kcal/mol lower in energy than the Oxetane-NH_4^+ complex and by 8.7 kcal/mol than the theitane-NH_4^+ complex. This result indicates that O4b is more basic than S4b. Rearrangement by protonation, thus, is more favorable for Oxetane than for thietane, consistent with the known increased stability of thietane.[32]

Discussion

Pyrimidine dimers are the major DNA photoproducts under solar irradiation in living organisms. Both residues in a bipyrimidine sequence can be photo-excited; whether these residues contribute equally to the dimerization reaction however is unclear. In the present contribution, we demonstrated that after replacing the 5'-residue by a *meta*-xylene, the resulting XpT dinucleotide after thymine photoexcitation readily supports 6-4PP photochemistry. We previously demonstrated that upon replacement of the 3'-pyrimidine with a toluenyl (To) or *meta*-xylenyl moiety (X) CPD analogs are formed as the dominant photoproducts under UV irradiation.[22] In the formation of SP in the TpT photoreaction,[11] the photo-excited 5'-thymine residue abstracts an hydrogen from the methyl group at the 3'-thymine.[13] Taken together, that *initial photoexcitation of only one of the two involving pyrimidine residues being sufficient to drive the dimerization to completion is likely a general model for all three types of pyrimidine photoreactions*.

Furthermore, our present results have indicated that the 6-4PP formation is triggered by the acceptance of an electron by the "driver" 3'-thymine residue from the "passenger" 5'-thymine residue. It is known that UV radiation may trigger intrastrand ET in DNA resulting in a pair of transient cation and anion radicals.[33] The transient radicals are indicated to be responsible for the formation of intrastrand crosslink products involving a 5-bromouracil or 5-bromocytosine and a neighboring purine residue where the pyrimidine anion eliminates a bromide before the resulting radical reacts with the purine cation to form the crosslinking bond.[34] Interbase ET in a DNA single strand $(dT)_{18}$ occurs on a sub-picosecond timescale, and the resulting radical cation and anion species decay on a 100 picosecond timescale.[35] Moreover, interbase ET in the TpT step may also play a key role in the 6-4PP formation.[7, 15a] However, experimental evidence for directionality of the ET is not easy to obtain due to the similar redox potential between the Ts making both 5′ to 3′ and 3′ to 5′ interbase ET possible. The substitution of a thymine by a xylene residue resulting in dinucleotides TpX and XpT removes this ambiguity. UV irradiation of the TpX complex in which 3′ to 5′ ET occurs only generates CPD.[22] In contrast, in the XpT studied here, 5′ to 3′ ET occurs that leads to 6-4PP formation and completely abolishes the CPD photochemistry. Therefore, our system provides experimental evidence that 6-4PP photochemistry is initiated by the 5′ to 3′ ET process.

Using extensive quantum-chemical calculations we demonstrated that the initially formed charge-separated radical pair further lowers its energy and subsequently recombines by formation of the interbase covalent bonds. In fact, the distinct chemical structure of the two bases in the 6-4PP and SP-like photoproducts is determined by interactions of the 3'-base radical anion and 5'-base radical cation. Under favorable reaction conditions in duplex DNA, the interbase ET leading to 6-4PP and SP formation may compete with other excitation-decay pathways such as interstrand proton transfer resulting from the H-bonding interactions between Watson-Crick base pairs implied by recent excited state dynamics studies in duplex DNA.[33a, d, e]

Formation of various pyrimidine dimers is long known to be wavelength-dependent.[36] The short-wavelength UVC light leads to the formation of all three dimers. In contrast, under the longer-wavelength UVB light, 6-4PP formation is drastically reduced; while no 6-4PP can be formed under UVA irradiation.[1e, 9b, 37] Moreover, in the presence of triplet-state photosensitizers, both CPD and SP can be formed under UVA light; again no 6-4PP was detected.[12] It is therefore concluded that both singlet and triplet excited states can contribute to CPD and SP formation, while only singlet excited states are responsible for 6-4PP formation.[1d] Our calculations demonstrated that photochemical formation of interbase covalent bonds proceeding through the triplet and singlet routes may differ drastically in timescales, formation of intermediates and quantum yields. In particular, the quantum yield of the triplet route might be rather low because radical recombination before the C6a−O4a bond is formed leads to the recovery of the initial dinucleotide rather than to 6-4PP formation.

The oxetane intermediate is widely accepted as an intermediate preceding 6-4PP formation when the 3'-residue is a thymidine. By replacing the 3'-thymidine with a 4-thiothymidine in a dinucleotide context,[32] a semi-stable thietane species was generated under 360 nm UVA irradiation, which was under equilibrium with the corresponding 6-4PP analog. Given the similar structures between thietane and oxetane, it is reasonable that the thymidine photoreaction may as well proceed via an

oxetane intermediate to form 6-4PP. However, although oxetane intermediates via the Paternò-Büchi reaction are relatively common in organic photochemistry,[38] the lifetime of the oxetane intermediate generated during the 6-4PP formation is estimated to be ~ 4 milliseconds in the dinucleotide TpT photoreaction.[39] Our calculations indicate that *oxetane rearrangement to 6-4PP may involve protonation of the basic O4b position*. The protonation mechanism also offers an explanation for the higher stability of the thietane intermediate,[32] as the S4b position is less basic as compared to O4b. In line with our finding, an equilibrium between thietane and the 6-4PP isomer is known to be pH dependent.[32] Increasing pH may lead to N3b deprotonation, and therefore, complete displacement of the equilibrium toward the 6-4PP isomer, as observed for thietane.[32] Under pH conditions when N3bH is protonated, protonation/deprotonation of S4b results in an equilibrium between the thietane and 6-4PP isomers.[32] As our calculations suggest, protonation of O4b is more energetically favorable as compared to S4b, hence the equilibrium should be completely displaced to the protonated 6-4PP form in the case of oxetane, possibly explaining the fact that oxetane is not observed upon XpT irradiation. Nonetheless, deprotonation of O4b orchestrated with protonation of N3b by a photolyase enzyme has been tentatively proposed as a crucial step in the reversal of 6-4PP to oxetane and then to the intact nucleobases in some organisms.[40]

The repair of CPD/6-4PP by a respective photolyase is initiated by intermolecular ET resulting in a radical anion specie of the photoproduct.[40-41] It is thus of interest for us to compare the ET-mediated 6-4PP formation with the photolyase mediated pyrimidine-dimer repair. Photolyases adopt a flavin adenine dinucleotide cofactor that absorbs visible light and in the excited state pumps an electron into the 5'-residue of the CPD/6-4PP to convert it to an anion radical to initiate the dimer repair process.[40-42] Such a trend is inverted as compared to the 6-4PP formation shown in our XpT system here, in which the transient anion radical is located at the 3'-residue. Apparently, interbase bond formation and cleavage in 6-4PP may be triggered by electron donation to the different bases, 3' and 5' respectively. The transient 5'-anion radical formed in photolyase then undergoes the repair possibly by the formation of a transient oxetane intermediate at the enzyme active site[42-43] or by direct transfer of the O4bH group from the 5'-base to the 3'-base.[44]

Although our conclusions on 6-4PP mechanism are reached by using a xylene as a thymine analog, the results are certainly consistent with the 6-4PP photochemistry of natural pyrimidine residues. The key structures and energy trends controlling this photochemistry identified in our present computational study for the XpT dinucleotide are similar to those found previously for the TpT and TpC dinucleotides.[7, 14, 29-30]

Besides the generation of 6-4PP analog **1**, the formation of SP analog **2** in the XpT photoreaction is also significant to the understanding of pyrimidine photochemistry triggered by interbase ET. In this case, the C4b−O4b group of the 3′-thymine radical anion abstracts a proton form the methyl-group of the 5'-radical cation after electron donation. Subsequent radical recombination results in the formation of the C5a−C4b bond. Such a reaction mechanism may be similar to the formation of SP, where the C5=C6 bond at the excited 5′-thymine abstracts an H atom from the 3′-thymine residue before the resulting thymine allyl radical recombines with the C5 radical.[11] The so-formed Spore-Int intermediate, likewise Oxetane, does not accumulate and instead undergoes elimination of a water molecule to form photoproduct **2.** As we did not detect any stable reaction intermediate during the formation of **2**, it is probable that the Spore-Int intermediate, like Oxetane, is unstable towards protonation or water-assisted proton transfer. *Therefore, besides 6-4PP formation, our work here presents the first indication that ET between the two bases may be responsible for SP-type photochemistry*.

Interestingly, a photoproduct with a structure similar to **2** was also observed in the previous thymidylyl (3'-5')-4-thiothymidine photoreaction,[32] suggesting that the photo-excited 4-thiothymidine can also initiate the H-atom abstraction process from the methyl group of the 5'-thymine base. The fact that both XpT and thymidylyl (3'-5')-4-thiothymidine exhibit the SP-type photoreaction indicates that such a route may be a general mechanism in pyrimidine photochemistry. As dinucleotides generally exhibit similar stacking patterns with those steps in the duplex DNA, the observed formation of **2** raises an interesting question: Can this C4=O mediated SP-type photoreaction occur in a duplex oligonucleotide and in the genomic DNA of solar radiated cells?

Conclusions

In summary, we report the synthesis and photoreaction of a dinucleotide analog XpT, in which a photo-inert xylene residue is incorporated into the 5'-end as a steric analog of thymine and as an electron donor. Irradiation of the XpT dinucleotide under 254 nm UV light generates two major photoproducts: a pyrimidine (6-4) pyrimidone (6-4PP) analog and an analog of the so-called spore photoproduct (SP). Both products are formed via reaction at the C4=O of the photo-excited T, which, together with our previous studies on the formation of cyclobutane pyrimidine dimers (CPDs) and SP, indicates that excitation of a single “driver” residue is sufficient to trigger pyrimidine dimerization and this is likely a general model for all three types of pyrimidine photoreactions.

Our observations that the photoreaction of TpX does not yield any 6-4PP but that of XpT does indicate that formation of 6-4PP and SP requires 5′ to 3′ interbase ET (but not 3′ to 5′) as the critical step for both detected photoreactions. Using quantum-chemical calculations we demonstrated that indeed 5′ to 3′ interbase ET catalyzes the formation of the interbase covalent bonds in 6-4PP and SP. The photochemistry of the XpT dinucleotide analog thus provides the first unambiguous evidence that the 6-4PP formation is triggered by the acceptance of an electron by the “driver” 3'-thymine residue from the “passenger” 5'-thymine residue. Our work also provides the first indication that interbase ET may be responsible for triggering the SP-type photochemistry.

Our quantum-chemical calculations imply that opening of the oxetane ring may be facilitated by protonation of the O4b atom.

Toluene-based thymine analogs like xylene are proved to exhibit similar base-stacking interactions as a thymine. Further, our calculations identified structures controlling the XpT photochemistry that are rather similar to those found in the previous computational studies of the TpT and TpC dinucleotides. Hence, even though our conclusion on the role of 5′ to 3′ ET in pyrimidine dimerization is reached by using a xylene residue, the result is certainly consistent with the photochemistry of natural pyrimidine residues.

Experimental and Computational Section

Materials and general methods. All solvents and chemicals were of analytical grade and purchased from Sigma-Aldrich, Fisher, or VWR and used without further purification. The 254 nm UV irradiation was provided by a Spectroline germicidal UV sterilizing lamp XX-15G, which generates unfiltered UV light ranging from 180 nm to 280 nm and peaking at 254 nm. The 302 nm UV irradiation is provided by using a UVP XX-15M lamp, which generates unfiltered UV light ranging from 280 nm to 360 nm and peaking at 302 nm. NMR spectra were obtained via a Bruker 500 MHz Fourier transform NMR spectrometer. Mass spectrometric (MS) analyses were conducted via electrospray ionization (ESI) employing an Agilent 1100-6130A LC-MS with an ion-trap mass analyzer. High resolution MS and tandem mass spectrometry (MS/MS) analyses were performed using an Agilent 1200-6520 capillary LC-Q-TOF MS spectrometer. The MS data were acquired via the "Agilent MassHunter Workstation Data Acquisition (B.03.00)" software and analyzed via "Qualitative Analysis of MassHunter Acquisition Data (B.03.00)" software.

Synthesis of XpT. The thymidine isostere (X) containing a *meta*-xylene moiety was synthesized via a procedure reported previously by our group.[22] We then linked X to a thymidine residue through a phosphodiester moiety, obtaining the dinucleotide analog XpT (Scheme 2). Detailed synthesis and characterization of XpT can be found in the supporting information.

XpT photoreaction under 254 nm UV light. XpT was dissolved in 200 µL 10 mM pH 7.4 phosphate buffer to a final concentration of 2 mM. The resulting solution was then irradiated under unfiltered 254 nm UVC light on ice with the samples ~ 5 cm away from the lamp. At various times, 10 µL of the solution was aliquoted out for later HPLC analysis. Two major products **1** and **2** were observed when monitored at 260 and 320 nm by a UV director on HPLC.

XpT photoreaction under 302 nm UV light. The XpT was repeated under unfiltered 302 nm UVB light under identical conditions and the reaction products analyzed by HPLC. No products were observed.

Quantum yield determination. The quantum yields for products **1** and **2** were determined using a modified literature method.[45] Briefly, 2 ml of aqueous solution containing equal amount of dinucleotides TpT and XpT was placed in a small beaker and stirred continuously under the 254 nm UV light. After 10 minutes, the photoreactions were stopped and the products formed in these two reactions analysed by HPLC.[23] Under the similar reaction conditions, the quantum yield of 6-4PP TpT was determined as $(2.0 \pm 0.7) \times 10^{-3}$.[45] The quantum yields of **1** and **2** were determined as $(3.2 \pm 1.0) \times 10^{-4}$ and $(3.3 \pm 1.0) \times 10^{-4}$, respectively by comparing with that of 6-4PP TpT.

Preparation of XpT photoproducts in a large scale. To characterize the XpT photoproducts **1** and **2**, the reaction was repeated at a bigger scale. Briefly, 200 mg XpT was dissolved in 10 mM pH 7.4 phosphate buffer to a final concentration of 2 mM. The solution was placed on ice and the photoreaction was allowed to proceed for 2 hours. The products were purified by semi-preparative HPLC. **1** and **2** were collected, concentrated, and desalted by reinjection into HPLC and using H_2O and acetonitrile as mobile phases. The reaction afforded 2.4 mg **1** (1.2 % yield) and 5.1 mg **2** (2.6% yield) as white solids. The desalted products were then characterized by MS/MS and NMR spectroscopy.

HPLC analyses. HPLC analyses were performed at room temperature using a Waters (Milford, MA) breeze HPLC system with a 2489 UV/Visible detector and a Waters XBridge OST C18 column (2.5 µm, 4.6×50 mm). The reaction was monitored at 230 nm and 320 nm simultaneously. The HPLC column was first equilibrated with 99% mobile phase A (10 mM ammonium acetate aqueous solution, pH 7.0) and 1% of mobile phase B (acetonitrile). Compounds were eluted with an ascending gradient (1% to 20% in 30 min) of mobile phase B at a flow rate of 1 mL/min. Under this gradient, **1** was eluted at 8.3 min, **2** at 14.2 min, and XpT at 23.9 min.

Product purification by semi-preparative HPLC was also performed at room temperature with the same Waters HPLC system and a Waters XBridge™ OST C18 column (2.5 µm particle size, 10 × 50 mm). 10 mM pH 7.0 ammonium acetate in H_2O was used as mobile phase A and acetonitrile was used as mobile phase B. The column was equilibrated with 99% mobile phase A and 1% of mobile phase B. Compounds were eluted with an ascending gradient (1 - 5% in 20 min followed by 5 - 20% in another 20 min) of mobile phase B at a flow rate of 4.73 mL/min. Under this gradient, **1** was eluted at 12.1 min, **2** at 26.0 min, and XpT at 34.0 min.

Product analyses via tandem mass spectrometry (MS/MS). MS/MS analyses of the reaction products were conducted using an Agilent 6520 Accurate Mass Q-TOF LC/MS spectrometer with an Agilent ZORBAX Bonus-RP column (5 µm particle size, 250 × 4.6 mm i.d.). The column was equilibrated in mobile phase A (5 mM ammonium acetate in 99% water and 1% acetonitrile, pH 6.5), while compounds were eluted with an ascending gradient (0 - 15% in 30 min followed by 15 - 35% in 20 min) of mobile phase B (5 mM ammonium acetate in 50% methanol and 50% acetonitrile) at a flow rate of 0.5 mL/min. Under this gradient, **1** was eluted at 24.1 min, **2** at 37.2 min, and XpT at 42.0 min. Mass spectrometry was conducted under both positive and negative ion modes.

Ground- and excited-state calculations. Quantum-chemical calculations were performed for XpT and TpT dinucleotides with deprotonated phosphate (total charge −1). Geometry optimization was performed with density-functional theory methods with the B3LYP functional[46] supplemented by the D3 correction for dispersion energy.[47] In the ground state and lowest-energy triplet state, geometries of the stationary points (minima and saddle points) were first optimized and then confirmed by the normal vibrational mode analysis in harmonic approximation. From each computed saddle point, intrinsic reaction coordinates (*irc*) were computed in both directions following the imaginary-frequency mode. The *irc* calculations for the formation of **1** and **2** converged to similar structures of the ground-state and triplet-state reactant corresponding to a stacked conformation of XpT. In the first excited singlet state, geometry optimization was performed with the TD-B3LYP-D3 method, starting from the ground-state stacked conformation of XpT. The relaxed energy scans for the C5a-O4b and Ha-O4b bond formation were computed in the first excited state corresponding to the charge-separated radical-pair state with the TD-B3LYP-D3 method.

Single-point excited-state calculations were carried out with the multireference multiconfigurational second-order perturbation theory methods XMCQDPT2[48] with the zeroth-order CASSCF wave function. The XMCQDPT2 calculations were combined with the equilibrium dielectric polarizable continuum model (DPCM)[49] of water. The CASSCF active space was composed from the frontier molecular orbitals (MOs) localized on xylene and thymine rings (Supporting Information Figures S8 and S9). CASSCF energy averaging over all included states (singlet states or both singlet and triplet states) with equal weights was performed. The XMCQDPT2 calculations were carried out with intruder-state avoidance energy shift (edshft) of 0.02 hartree. Optimization of the conical intersections geometries[50] was carried out with the SA2-CASSCF(12,12) method (12 electrons in 12 MOs with energy-averaging over 2 states). The Mulliken atomic charges derived from the zeroth-order QDPT electron densities were used to determine the net charges on the bases. In the XpT and TpT dinucleotides, electronic coupling and maximum ET rate constants were evaluated using the generalized Mulliken-Hush method[27] and semi-classical Marcus theory.[28]

Thymine and xylene methylated at position N1/C1 were used in calculations of the redox energies. Geometry optimization and energy calculations were performed with the (U)B3LYP-D3-PCM method. Electron affinities and ionization potentials were computed as adiabatic energy differences between the radical ion and neutral closed-shell (*CS*) or triplet species. A positive electron affinity corresponds to the radical anion being lower in energy than the neutral *CS* or triplet species. The positive ionization potential corresponds to the radical cation being higher in energy than the neutral *CS* or triplet species.

The standard 6-31G** basis set was used throughout. All calculations were performed with Firefly 8.1 software[51] which is partially based on the US GAMESS code.[52]

Acknowledgements

The authors thank the National Science Foundation (CHE 1454184 to L.L) and the Minerva Program of the Max-Planck Society (to T.D.) for financial support. The NMR and MS facilities at IUPUI are supported by National Science Foundation MRI grants CHE-0619254 and DBI-0821661, respectively. T.D. and E.M. are very grateful to Chris Roome (MPImF Heidelberg) for excellent support of high performance computing.

Keywords: thymine dimer • DNA photochemistry • electron transfer • pyrimidine (6-4) pyrimidone photoproduct • spore photoproduct

[1] a) J. Cadet and P. Vigny in *Photochemistry and the nucleic acids*, *Vol. 1* (Ed. H. Morrison), Wiley, New York, **1990**, pp. 1-272; b) J. Cadet, E. Sage and T. Douki, *Mutat. Res-Fund. Mol. M.* **2005**, *571*, 3-17; c) J. Cadet, A. Grand and T. Douki in *Solar UV radiation-induced DNA bipyrimidine photoproducts: formation and mechanistic insights*, *Vol. 356* Eds.: M. Barbatti, A. C. Borin and S. Ullrich), Springer International Publishing, **2015**, pp. 249-275; d) J. Cadet, S. Mouret, J. L. Ravanat and T. Douki, *Photochem. Photobiol.* **2012**, *88*, 1048-1065; e) J. Cadet, S. Courdavault, J. L. Ravanat and T. Douki, *Pure Appl. Chem.* **2005**, *77*, 947-961.
[2] R. E. Rycyna and J. L. Alderfer, *Nucleic Acids Res.* **1985**, *13*, 5949-5963.
[3] J. S. Taylor and M. P. Cohrs, *J. Am. Chem. Soc.* **1987**, *109*, 2834-2835.
[4] a) C. l. Desnous, D. Guillaume and P. Clivio, *Chem. Rev.* **2010**, *110*, 1213-1232; b) P. Setlow, *Trends Microbiol.* **2007**, *15*, 172-180; c) P. Setlow, *J. Appl. Microbiol.* **2006**, *101*, 514-525; d) P. Setlow and L. Li, *Photochem. Photobiol.* **2015**, *91*, 1263-1290.
[5] a) D. E. Brash, *Photochem. Photobiol.* **2015**, *91*, 15-26; b) E. D. Pleasance, R. K. Cheetham, P. J. Stephens, D. J. McBride, S. J. Humphray, C. D. Greenman, I. Varela, M.-L. Lin, G. R. Ordonez, G. R. Bignell, K. Ye, J. Alipaz, M. J. Bauer, D. Beare, A. Butler, R. J. Carter, L. Chen, A. J. Cox, S. Edkins, P. I. Kokko-Gonzales, N. A. Gormley, R. J. Grocock, C. D. Haudenschild, M. M. Hims, T. James, M. Jia, Z. Kingsbury, C. Leroy, J. Marshall, A. Menzies, L. J. Mudie, Z. Ning, T. Royce, O. B. Schulz-Trieglaff, A. Spiridou, L. A. Stebbings, L. Szajkowski, J. Teague, D. Williamson, L. Chin, M. T. Ross, P. J. Campbell, D. R. Bentley, P. A. Futreal and M. R. Stratton, *Nature* **2010**, *463*, 191-196.
[6] a) H. E. Johns, M. Delbruck and S. A. Rapaport, *J. Mol. Biol.* **1962**, *4*, 104-114; b) A. J. Varghese and S. Y. Wang, *Science* **1967**, *156*, 955-957; c) J. E. Donnellan, Jr. and R. B. Setlow, *Science* **1965**, *149*, 308-310.
[7] A. Banyasz, T. Douki, R. Improta, T. Gustavsson, D. Onidas, I. Vaya, M. Perron and D. Markovitsi, *J. Am. Chem. Soc.* **2012**, *134*, 14834-14845.
[8] W. J. Schreier, P. Gilch and W. Zinth, *Annu. Rev. Phys. Chem.* **2015**, *66*, 497-519.
[9] a) A. Banyasz, I. Vaya, P. Changenet-Barret, T. Gustavsson, T. Douki and D. Markovitsi, *J. Am. Chem. Soc.* **2011**, *133*, 5163-5165; b) S. Mouret, C. Baudouin, M. Charveron, A. Favier, J. Cadet and T. Douki, *Proc. Natl. Acad. Sci.* **2006**, *103*, 13765-13770.
[10] S. Premi, S. Wallisch, C. M. Mano, A. B. Weiner, A. Bacchiocchi, K. Wakamatsu, E. J. Bechara, R. Halaban, T. Douki and D. E. Brash, *Science* **2015**, *347*, 842-847.
[11] a) D. M. Ames, G. Lin, Y. Jian, J. Cadet and L. Li, *J. Org. Chem.* **2014**, *79*, 4843-4851; b) G. Lin and L. Li, *Angew. Chem. Int. Ed.* **2010**, *49*, 9926-9929; c) E. C. Hayes, Y. Jian, L. Li and S. Stoll, *J. Phys. Chem. B* **2016**, *120*, 10923-10931; d) Y. Jian, D. M. Ames, H. Ouyang and L. Li, *Org. Lett.* **2015**, *17*, 824-827.
[12] A. Moysan, A. Viari, P. Vigny, L. Voituriez, J. Cadet, E. Moustacchi and E. Sage, *Biochemistry* **1991**, *30*, 7080-7088.
[13] Q. Du, H. Zhao, D. Song, K. Liu and H. Su, *J. Phys. Chem. B* **2012**, *116*, 11117-11123.
[14] A. Giussani, L. Serrano-Andrés, M. Merchán, D. Roca-Sanjuán and M. Garavelli, *J. Phys. Chem. B* **2013**, *117*, 1999-2004.
[15] a) R. Improta, *J. Phys. Chem. B* **2012**, *116*, 14261-14274; b) L. Esposito, A. Banyasz, T. Douki, M. Perron, D. Markovitsi and R. Improta, *J. Am. Chem. Soc.* **2014**, *136*, 10838-10841; c) A. Banyasz, L. Esposito, T. Douki, M. Perron, C. Lepori, R. Improta and D. Markovitsi, *J. Phys. Chem. B* **2016**, *120*, 4232-4242; d) D. Markovitsi, *Photochem. Photobiol.* **2016**, *92*, 45-51.
[16] a) W. J. Schreier, T. E. Schrader, F. O. Koller, P. Gilch, C. E. Crespo-Hernandez, V. N. Swaminathan, T. Carell, W. Zinth and B. Kohler, *Science* **2007**, *315*, 625-629; b) T. Ostrowski, J.-C. Maurizot, M.-T. Adeline, J.-L. Fourrey and P. Clivio, *J. Org. Chem.* **2003**, *68*, 6502-6510; c) C. Moriou, M. Thomas, M.-T. Adeline, M.-T. Martin, A. Chiaroni, S. Pochet, J.-L. Fourrey, A. Favre and P. Clivio, *J. Org. Chem.* **2007**, *72*, 43-50.
[17] a) B. A. Schweitzer and E. T. Kool, *J. Am. Chem. Soc.* **1995**, *117*, 1863-1872; b) K. M. Guckian and E. T. Kool, *Angew. Chem. Int. Ed.* **1997**, *36*, 2825-2828; c) E. T. Kool and H. O. Sintim, *Chem. Commun.* **2006**, 3665-3675.
[18] a) S. Moran, R. X.-F. Ren and E. T. Kool, *Proc. Natl. Acad. Sci.* **1997**, *94*, 10506-10511; b) S. Moran, R. X. F. Ren, S. Rumney and E. T. Kool, *J. Am. Chem. Soc.* **1997**, *119*, 2056-2057; c) E. T. Kool, *Annu. Rev. Biochem.* **2002**, *71*, 191-219.
[19] a) K. M. Guckian, T. R. Krugh and E. T. Kool, *Nat. Struct. Biol.* **1998**, *5*, 954-959; b) K. M. Guckian, T. R. Krugh and E. T. Kool, *J. Am. Chem. Soc.* **2000**, *122*, 6841-6847.
[20] P. S. Pallan and M. Egli, *J. Am. Chem. Soc.* **2009**, *131*, 12548-12549.
[21] P. R. Callis, *Annu. Rev. Phys. Chem.* **1983**, *34*, 329-357.
[22] a) D. Liu and L. Li, *RSC Adv.* **2013**, *3*, 19545-19550; b) D. Liu, Y. Zhou, J. Pu and L. Li, *Chem. Eur. J.* **2012**, *18*, 7823-7833.
[23] *See supporting information*.
[24] S. Iwai, M. Shimizu, H. Kamiya and E. Ohtsuka, *J. Am. Chem. Soc.* **1996**, *118*, 7642-7643.
[25] G. Lin, Y. Jian, H. Ouyang and L. Li, *Org. Lett.* **2014**, *16*, 5076-5079.
[26] a) J.-S. Taylor, D. S. Garrett and M. J. Wang, *Biopolymers* **1988**, *27*, 1571-1593; b) L. S. Kan, L. Voituriez and J. Cadet, *Biochemistry* **1988**, *27*, 5796-5803.
[27] C.-P. Hsu, *Acc. Chem. Res.* **2009**, *42*, 509-518.
[28] R. A. Marcus and N. Sutin, *BBA Rev. Bioenergetics* **1985**, *811*, 265-322.
[29] a) Z. b. Yang, R. b. Zhang and L. A. Eriksson, *Phys. Chem. Chem. Phys.* **2011**, *13*, 8961-8966; b) Z. b. Yang, L. A. Eriksson and R. b. Zhang, *J. Phys. Chem. B* **2011**, *115*, 9681-9686.
[30] L. Blancafort and A. Migani, *J. Am. Chem. Soc.* **2007**, *129*, 14540-14541.
[31] V. Labet, N. Jorge, C. Morell, T. Douki, A. Grand, J. Cadet and L. A. Eriksson, *Photochem. Photobiol. Sci.* **2013**, *12*, 1509-1516.
[32] P. Clivio, J. L. Fourrey, J. Gasche and A. Favre, *J. Am. Chem. Soc.* **1991**, *113*, 5481-5483.
[33] a) Y. Zhang, K. de La Harpe, A. A. Beckstead, L. Martínez-Fernández, R. Improta and B. Kohler, *J. Phys. Chem. Lett.* **2016**, *7*, 950-954; b) Y. Zhang, J. Dood, A. A. Beckstead, X.-B. Li, K. V. Nguyen, C. J. Burrows, R. Improta and B. Kohler, *Proc. Natl. Acad. Sci.* **2014**, *111*, 11612-11617; c) D. B. Bucher, B.

M. Pilles, T. Carell and W. Zinth, *Proc. Natl. Acad. Sci.* **2014**, *111*, 4369-4374; d) Y. Zhang, K. de La Harpe, A. A. Beckstead, R. Improta and B. Kohler, *J. Am. Chem. Soc.* **2015**, *137*, 7059-7062; e) Y. Zhang, X. B. Li, A. M. Fleming, J. Dood, A. A. Beckstead, A. M. Orendt, C. J. Burrows and B. Kohler, *J. Am. Chem. Soc.* **2016**, *138*, 7395-7401.
[34] a) Y. Zeng and Y. Wang, *Nucleic Acids Res.* **2006**, *34*, 6521-6529; b) H. Hong and Y. Wang, *J. Am. Chem. Soc.* **2005**, *127*, 13969-13977.
[35] B. M. Pilles, D. B. Bucher, L. Liu, P. Gilch, W. Zinth and W. J. Schreier, *Chem. Commun.* **2014**, *50*, 15623-15626.
[36] T. Matsunaga, K. Hieda and O. Nikaido, *Photochem. Photobiol.* **1991**, *54*, 403-410.
[37] T. Douki, A. Reynaud-Angelin, J. Cadet and E. Sage, *Biochemistry* **2003**, *42*, 9221-9226.
[38] M. Auria and R. Racioppi, *Molecules* **2013**, *18*, 11384.
[39] S. Marguet and D. Markovitsi, *J. Am. Chem. Soc.* **2005**, *127*, 5780-5781.
[40] A. Sancar, *Chem. Rev.* **2003**, *103*, 2203-2238.
[41] A. Sancar, *J. Biol. Chem.* **2008**, *283*, 32153-32157.
[42] Y. Wang, P. P. Gaspar and J.-S. Taylor, *J. Am. Chem. Soc.* **2000**, *122*, 5510-5519.
[43] a) X. Zhao, J. Liu, D. S. Hsu, S. Zhao, J.-S. Taylor and A. Sancar, *J. Biol. Chem.* **1997**, *272*, 32580-32590; b) S. T. Kim, K. Malhotra, C. A. Smith, J. S. Taylor and A. Sancar, *J. Biol. Chem.* **1994**, *269*, 8535-8540.
[44] a) T. Domratcheva and I. Schlichting, *J. Am. Chem. Soc.* **2009**, *131*, 17793-17799; b) T. Domratcheva, *J. Am. Chem. Soc.* **2011**, *133*, 18172-18182.
[45] Y. K. Law, J. Azadi, C. E. Crespo-Hernández, E. Olmon and B. Kohler, *Biophys. J.* **2008**, *94*, 3590-3600.
[46] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648-5652.
[47] S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.
[48] A. A. Granovsky, *J. Chem. Phys.* **2011**, *134*, 214113.
[49] a) R. Cammi and J. Tomasi, *J. Comput. Chem.* **1995**, *16*, 1449-1458; b) B. Mennucci and J. Tomasi, *J. Chem. Phys.* **1997**, *106*, 5151-5158.
[50] A. A. Granovsky, *J. Chem. Phys.* **2015**, *143*, 231101.
[51] *Alex A. Granovsky, Firefly version 8, www http://classic.chem.msu.su/gran/firefly/index.html, access date March 2017.*
[52] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis and J. A. Montgomery, *J. Comput. Chem.* **1993**, *14*, 1347-1363.

Entry for the Table of Contents (Please choose one layout)

Layout 2:

FULL PAPER

Yajun Jian,[† [a, b]] Egle Maximowitsch,[† [c]] Degang Liu,[[a]] Surya Adhikari,[[a]] Lei Li,[[a,d]] and Tatiana Domratcheva*[[c]]*

Indications of 5′ to 3′ Interbase Electron Transfer as the First Step of Pyrimidine Dimer Formation Probed by a Dinucleotide Analog