SN- and NS-puckered sugar conformers are precursors of the (6–4) photoproduct in thymine dinucleotide

Jouda Jakhlal, Clément Denhez, Stéphanie Coantic-Castex, Agathe Martinez, Dominique Harakat, Thierry Douki, Dominique Guillaume, Pascale Clivio

▶ **To cite this version:**

Jouda Jakhlal, Clément Denhez, Stéphanie Coantic-Castex, Agathe Martinez, Dominique Harakat, et al.. SN- and NS-puckered sugar conformers are precursors of the (6–4) photoproduct in thymine dinucleotide. Organic & Biomolecular Chemistry, 2022, 20 (11), pp.2300-2307. 10.1039/d2ob00044j . hal-03850053

HAL Id: hal-03850053
https://hal.science/hal-03850053

Submitted on 12 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SN- and NS-Puckered sugar conformers are precursors of the (6-4) photoproduct in thymine dinucleotide

Received 00th January 20xx,
Accepted 00th January 20xx

Jouda Jakhlal,[a,‡] Clément Denhez,[a,b] Stéphanie Coantic-Castex,[a] Agathe Martinez,[c] Dominique Harakat,[c] Thierry Douki,[d] Dominique Guillaume[a] and Pascale Clivio*[a]

Some amount of furanose in a southern conformation, possibly in both, but certainly in one of the two adjacent nucleotides of a dipyrimidine site, is necessary for (6-4) photoproduct formation in oligonucleotides. To explore the necessity, role, and most favorable location of each South sugar conformer on the formation of the (6-4) adduct in the thymine dinucleotide TpT, the photochemical behavior of two synthetic analogues, in which the South sugar conformation is prohibited for one of their two sugars, has been examined. Herein, we experimentally demonstrate that the presence of one sugar presenting some amount of South puckering, at any of the extremities, is sufficient to trigger (6-4) adduct formation. Nonetheless, the photochemical behavior of the dinucleotide with a South-puckered conformation at the 5'-end, mimics more closely that of TpT. In addition, using the 5' North 3' South-dilocked dinucleotide, we demonstrate that the flexibility of the South pucker at the 3'-end has little influence on the (6-4) adduct formation.

Introduction

Two distinct photocrosslinking reactions between adjacent pyrimidine nucleobases constitute the major source of damage in solar UV-exposed DNA; giving rise to cyclobutane pyrimidine dimers (CPDs) and (6-4) photoproducts ((6-4) PPs) (Fig. 1).[1] DNA PPs are critical actors in the pathogenesis of UV-induced skin cancer, eye diseases and photoaging. [2] DNA PPs are also essential for the equilibrium of ecosystems,[3] and can trigger the production of important secondary metabolites in plants.[4] In UV-exposed DNA, compared to CPDs, the occurrence of (6-4) PPs is only about one-third and their repair is much faster.[5,6] However, their biological importance should not be underrated since, if inefficiently repaired, they are highly cytotoxic and mutagenic.[7] More importantly, (6-4) PPs (and not CPDs) have very recently been shown to activate the ATR-Chk1 pathway, a master process in DNA damage response.[8] Numerous experimental and theoretical studies have focussed on the photochemical formation mechanism of CPD and (6-4) DNA PPs.[9] It is currently accepted that CPD principally derives from an excited singlet state and its formation (~ 1 ps) is controlled by the ground state conformation of the two reacting nucleobases.[10]

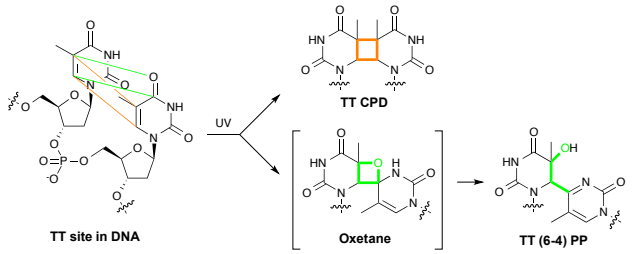


Fig. 1 Structure of the two major photoproducts formed at dithymine sites in DNA.

Conversely, the formation mechanism of (6-4) PP is not fully understood, yet. It likely involves the initial formation of a transient four membered ring (oxetane or azetidine) intermediate whose putative formation has been originally characterized in the 4-thiothymine series.[11] Very recently, a precursor of the thietane intermediate has been detected and shown to derive from the triplet state of 4-thiothymine in ~ 3 ps after UV excitation and ultrafast intersystem crossing.[12] Hypothesizing a similar pathway and time scale in the three four-membered ring series, the ground state conformation of the two pyrimidine moieties, at the time of UV excitation, must control the formation of the oxetane (Fig. 1) and azetidine intermediates. Accordingly, numerous studies have shown that the conformation of naked or protein-associated DNA has an incidence on the (6-4) PP yield.[13,14] The most recurrent, but rather imprecise conformational factor associated with (6-4) PP formation is the requirement of a certain degree of flexibility of the sugar phosphate DNA backbone. Experimental[15] and theoretical[16] studies have also led to identify the nucleotide sugar conformation as a major determinant in the production of (6-4) adduct at the single-strand di- and oligonucleotide level. However, further

[a] *Université de Reims Champagne Ardenne, Institut de Chimie Moléculaire de Reims, CNRS UMR 7312, UFR de Pharmacie, 51100 Reims, France.*
[b] *MaSCA, P3M, UFR des Sciences Exactes et Naturelles, 51100 Reims, France.*
[c] *Université de Reims Champagne Ardenne, Institut de Chimie Moléculaire de Reims, CNRS UMR 7312, UFR des Sciences Exactes et Naturelles, 51100 Reims, France.*
[d] *Université Grenoble Alpes, CEA, CNRS, IRIG, SyMMES, 38000 Grenoble, France.*
‡ *Present address: Evotec SAS, Campus Curie, 31036 Toulouse Cedex, France.*
†Electronic Supplementary Information (ESI) available: See
DOI: 10.1039/x0xx00000x

experimental studies are still required to provide an in-depth understanding of the role played by the sugar conformation in (6-4) PP production. Upon irradiation at 254 nm, the dinucleotide TpT affords CPD and (6-4) PP. It is therefore a particularly efficient tool to decrypt the photoreactivity of the dithymine site within DNA.[17,18] Advantageously, this model also allows the study of the intrinsic participation of the sugar phosphate backbone on the formation of CPD and (6-4) PP out of the context of any base pairing, flanking sequences or protein interactions. Such information furnishes the initial basis for a better understanding of what happens next in DNA and DNA protein complexes. In addition, the inherent flexibility of TpT is welcomed since highly fluxional DNA structures allow the formation of (6-4) adducts.[13f]

The sugar residues of **1** are in a dynamic C3'-endo (North)/C2'-endo (South) equilibrium and exhibit predominantly South (S)-puckered conformers (X_S: 5'-end: 72%, 3'-end: 59 %).[19] Through chemical modification of the sugar residues of **1**, the North (N) population of conformers at the 5'- and 3'-end has been increased to 78 and 65%, respectively.[15a,b] As a consequence, (6-4) PP formation increased.[15a,b] However, when the amount of N population at both ends nears 95%, the efficiency of (6-4) PP production decreases to return to that of TpT.[15g] Finally, when the sugar residue of each extremity is entirely locked in the N conformation as in $T_{LN}pT_{LN}$ (**2**), (6-4) PP formation is hampered.[15d] This result has lent to support the idea that the presence of some amount of S sugar conformation is essential for the formation of (6-4) PP. Therefore, the NS, SN and SS conformations are potential appropriate candidates for (6-4) PP formation. In order to identify the location(s) of S puckered-sugar(s) suitable for (6-4) PP formation, here we report the photochemical behavior of the two hemilocked dinucleotides $T_{LN}pT$ (**3**) and TpT_{LN} (**4**) (Fig. 2),[20] in which one sugar residue is locked in the disfavorable N conformation whereas the other can freely populate N and S conformation(s). Our results are also discussed to the light of the NS-dilocked dinucleotide $T_{LN}pT_{LS}$ (**5**) (Fig. 2)[20] whose 3'-end is totally in a S-freezed conformation.

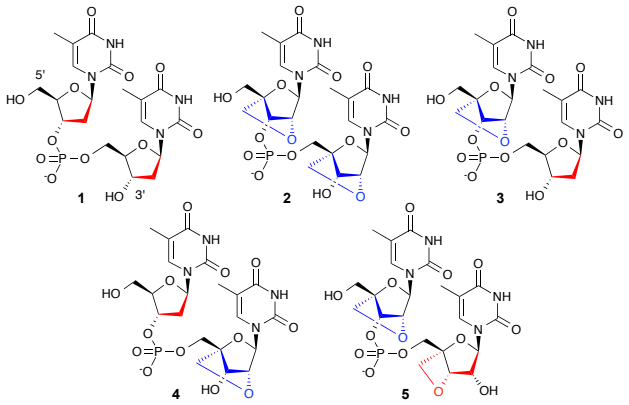


Fig. 2 Structures of the natural dinucleotide TpT (**1**), the NN-dilocked dinucleotide $T_{LN}pT_{LN}$ (**2**), the N-hemilocked dinucleotides $T_{LN}pT$ (**3**) and TpT_{LN} (**4**), and the NS-dilocked dinucleotide $T_{LN}pT_{LS}$ (**5**).

The known CD-derived stacking of **3**-**5**[20] will be important to consider with regard to the photochemical results.

Results and discussion

Synthesis of the dinucleotide analogues

The synthesis of dinucleotides $T_{LN}pT$ (**3**) and TpT_{LN} (**4**) has been previously reported.[20] The synthesis of $T_{LN}pT_{LS}$ (**5**) made use of the two known locked thymine nucleosides T_{LN} (**6**) and T_{LS} (**7**) previously prepared by the Vorbrüggen method.[21,22] Nucleoside **6** has also been synthesized from a locked uridine nucleoside through C5 methylation.[23] However, to our knowledge, neither **6** nor **7**, so far, have been directly prepared from the commercially available ribothymidine (5-methyluridine). Inspired by the reported divergent synthesis of 2'-*O*,4'-*C*- and 3'-*O*,4'-*C*-locked uridines from a uridine nucleoside,[24] we envisioned the synthesis of **6** and **7** from 2',3'-isopropylideneribothymidine (**8**)[25] (Scheme 1, conditions†). We only slightly modified the reported strategy[24] by using *o*-iodoxybenzoic acid (IBX) in acetonitrile[26] instead of the Moffatt conditions to oxidize **8** in **9** (92% yield**)**. The subsequent reaction conditions leading to **6** and **7** were essentially the same as those reported in the uridine series (Scheme 1).[24]

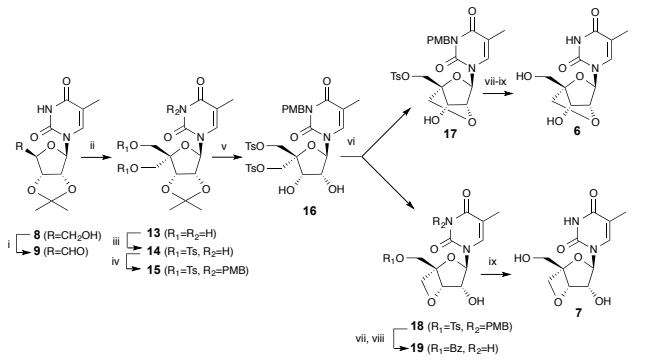
Nucleoside **6** afforded the known phosphoramidite **10** following literature procedures.[21b] Dimethoxytritylation of **7** afforded **11** (54% yield) that, upon a quantitative acetylation/detritylation procedure,[20] gave rise to the 2'-acetyl derivative **12**. Then phosphoramidite **10** and alcohol **12** were condensed in the presence of 5-(ethylthio)-1*H*-tetrazole and the resulting phosphite triester intermediate was oxidized and fully deprotected[20] affording dinucleotide **5** with 8 % yield (four steps) after HPLC purification (Scheme 2, conditions†).

Conformational analysis

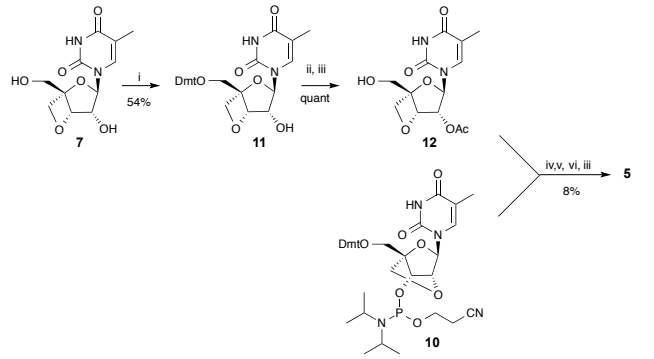
In nucleic acids, the conformation of the sugar is generally described by two parameters based on the concept of pseudorotation: the pseudorotational phase angle (*P*) that designates the out of plane atom(s) and the maximum puckering amplitude (ν_{max}) that specifies the out of plane extent. The solution-phase conformations adopted by the sugar residues of **3**-**5** were studied by ^{1}H NMR using experimental $^3J_{H,H}$ vicinal proton coupling constants at different temperatures (ESI Page S35 and Table S1†). Assuming a two state conformer equilibrium for the sugars of **3**-**5**, their pseudorotational parameters (P_1, ν_{1max}, P_2, ν_{2max}; respectively) as well as their fractional populations (X_1 and X_2, respectively) (Table 1) were determined using the Matlab Pseudorotation GUI program (ESI Pages S36-S71†).[27] DFT calculations were performed to obtain *Aj* and *Bj* coefficients required by the program to correlate external torsion angles (related to $^3J_{H,H}$) to the corresponding internal torsion angles (related to the sugar conformation).

NMR-based conformational analysis confirmed that N-locked sugar residues in **3**-**5** were exclusively in the N conformational domain with P_N and ν_{max} centered around 18° and 57°, respectively (entries 3, 6 and 7).

This journal is © The Royal Society of Chemistry 20xx



Scheme 1 Synthesis of **6** and **7** from 2',3'-isopropylidene-5-methyluridine (**8**). Reaction conditions: (i) IBX, CH_3CN; (ii) Formaldehyde, NaOH; (iii) TsCl, DMAP, CH_2Cl_2; (iv) PMBCl, K_2CO_3, DMF; (v) TFA/H_2O (9/1); (vi) NaH, DMF; (vii) CAN, H_2O/CH_3CN; (viii) Sodium benzoate, DMF; (ix) conc. NH_4OH.



Scheme 2 Synthesis of $T_{LN}pT_{LS}$ (**5**). [a]Reaction conditions: (i) DmtCl, pyridine; (ii) Ac_2O, pyridine; (iii) 80% aq AcOH; (iv) 5-(Ethylthio)-1*H*-tetrazole, CH_3CN; (v) I_2, THF/H_2O/2,6-lutidine; (vi) conc. NH_4OH.

Such results are fully in line with NMR/theoretical calculations and X-ray crystallographic analysis performed on the N-locked nucleoside **6**, free[28,29] or incorporated into oligonucleotides.[30,31] The large ν_{max} (entries 3, 6 and 7) resulted from the 2'-*O*,4'-*C*-methylene linkage constraint.

The N/S population equilibrium of the deoxyribose moieties in **3** and **4** did not substantially change compared to that in TpT (**1**) (entries 1, 2, 4 and 5).

Table 1 Pseudorotation Phase Angle (*P*, °), Maximum Puckering Amplitude (ν_{max}, °), and population of conformers (X, %) at 298 K for the sugar moieties of dinucleotides 1[19] and 3-5[a,b]

Entry	Compound	P_1	ν_{1max}	X_1	P_2	ν_{2max}	X_2	RMSd (Hz)
1	**1** (Tp-)	-19	46	28	141	26	72	
2	**1** (-pT)	18	35	41	149	33	59	
3	**3** (T_{LN}p-)	16.9	58.3	10.2	15.8	55.9	89.8	0.04
4	**3** (-pT)	8.0	46.5	38.6	114.5	26.4	61.4	0.11
5	**4** (Tp-)	3.5	36.0	36.6	143.1	25.7	63.4	0.09
6	**4** (-pT_{LN})	18.7	59.6	11.3	18.0	58.4	88.7	0.01
7	**5** (T_{LN}p-)	19.5	59.8	8.3	18.9	58.6	91.7	0.02
8	**5** (-pT_{LS})	131.6	31.9	49.3	153.6	27.5	50.7	0.02

[a]Calculated using Matlab pseudorotation GUI program. [b]*P* and X values of major S or N conformers are in red or blue, respectively.

The same trend was also observed for the puckering mode of the deoxyribose residues that remained in the same section of the pseudorotation wheel. The only difference was noticed for the S conformer at the 3'-end in **3** that lied in the C1'-exo domain (P_S = 114.5°) instead of the C2'-endo region in **1** (P_S = 141°) (entry 4). Therefore, the presence of the N-locked sugar at the 5'-end of **3** induced a slight conformational change of the flexible 3'-neighboring sugar moieties towards C1'-exo. A single N-locked nucleoside **6** incorporated in a DNA strand, when paired to a DNA or RNA strand, is known to shift the sugar conformation of the 3'-neighboring deoxyribonucleotide residue toward an N-type conformation.[30] In our single strand dinucleotide context, this effect is modest. This might indicate that, in the duplex series, the observed 3'-directional conformational steering consecutive to the presence of a single N-locked sugar residue, would not be due to backbone preorganization but could rather result from hybridization with the partner strand.

The 3'-*O*,-4'-*C* locked sugar in **5** was calculated (Matlab Pseudorotation) to be composed of two equal populations of S-type conformers: a C1'-exo conformer (*P* = 132°) and a C2'-endo conformer (*P* = 154°) with a ν_{max} of 32° and 27°, respectively (entry 8). Indeed, the restriction of the 3'-end sugar residue of **5** to the S conformational domain in solution is fully in accordance with the reported 94% C1'-exo-C2'-endo (*P* ≃ 144°; average between *P* = 126° (C1'-exo) and *P* = 162° (C2'-endo)) solution population of the sugar residue of nucleoside **7**.[22b,32] In addition, presence at the 3'-end sugar residue of **5** of a C1'-exo conformer was consistent with the X-ray structure of nucleoside **7** (*P* = 136° and ν_{max} = 32°) and PM3 calculations.[22] Regarding the probability of existence of a C2'-endo conformation, DFT calculations performed on a 3'-sugar residue model of **5** initially identified it as a minor conformer (10%, ESI Table S6†) but as a major conformer (55%, ESI Table S13†) after incorporation of an H-bonded water molecule as in the X-ray structure of **7**.

Photochemistry

The photochemical behavior of dinucleotides **3**-**5** was individually studied at 254 nm, in aqueous solution, and in the presence of **1** as an internal reference, as previously described.[15b-d,g,h] Photoproduct identification was performed on each crude photoreaction mixtures using a well-established HPLC/electrospray/tandem mass spectrometry (ES-MS/MS) method.[5,17] Inspection of chromatograms of the irradiated solutions of **3**-**5** in the presence of **1**, recorded using a photodiode array detector (multi wavelength UV detection, 200-400 nm), indicated the formation of only two photoproducts from each dinucleotide even though the possible formation of trace amount of photoproducts and/or degraded material, undetected under our experimental conditions, could not be excluded (ESI Fig. S44-46†). UV-induced (6-4) PPs from **3**-**5** were identified by their UV absorption near 330 nm and their main $[M-H-113]^-$ fragment ion on their MS/MS spectrum (*m/z* at 460 for the (6-4) PP of **3** and **4** and *m/z* at 488 for the (6-4) PP of **5**) (ESI Fig. S45-

49†).[17,33] The [M-H-113]$^-$ fragment ion would result from rearrangements of the dihydropyrimidine part of the (6-4) PP ultimately leading to opening and loss of a $C_5H_6NO_2$ fragment.[17,33] CPDs derived from dinucleotides **3** and **4** were identified by the absence of their UV absorption above 240 nm and their main [M-H-98]$^-$ ion at *m/z* 475 (ESI Fig. S47, S48†), corresponding to the loss of a didehydrated 2-deoxyribose from their 3'- and 5'-end, respectively (ESI Schemes S1, S2†).[17] In the crude irradiation mixture of $T_{LN}pT_{LS}$ (**5**), a PP assigned to a CPD from its UV spectrum exhibited a fragment ion at *m/z* 475 (ESI Fig. S49†). This latter was interpreted as a loss of a didehydrated ribose containing the 3',4'-OCH_2 linkage ([M-H-126]$^-$) (ESI Scheme S3†). Thus, the [M-H-126]$^-$ fragment ion likely originates from the 3' S-locked sugar residue of the CPD derived from **5**.

The time-dependent conversion of dinucleotides **3**-**5**, in the presence of **1**, to photoproducts upon 254 nm was followed by RP HPLC and the kinetic of the photochemical reactions expressed as the fractional amount (FA or relative chemical yield) of products formed[15c] vs the irradiation time (Fig. 3-5, ESI Page S80†). This indicated that **3**-**5** were able to qualitatively mimic the photochemical behavior of **1** though with a different photoproduct efficiency. The non linearity of the plot of **1**, **3**-**5** and of their corresponding CPD after t ~ 4 min is the consequence of the absorbance of CPD at 254 nm and its consecutive photoreversion.[18,15c] The linearity of the plots at t ≤ 4 min, due to negligible photoreversion, allowed to compare the photochemical efficiency between **1**, **3**-**5**. At t = 4 min, dinucleotide **3** was found to be the most photoreactive (FA = 63%), followed by **4** (FA = 69%) then **5** that is as photoreactive as **1** (FA = 72%). FA for (6-4) PP, whatever the dinucleotide, never exceeded that of **1** (FA = 3%) and was 2%, 0.8%, and 0.5% for **4**, **5**, and **3** respectively. Considering CPD formation, all three dinucleotides were more prone than **1** to form CPD (FA = 15%, 39%, 26%, and 20% for **1**, **3**, **4**, and **5**; respectively).

The efficiency of a photochemical process is also characterized by the quantum yield (Φ) that represents the amount of photoproduct formed per absorbed photon. In the case of the (6-4) adduct, the experimental Φ represents a global parameter that includes the light-induced oxetane formation and its rearrangement. For our study, to determine the relationship between chemical modification/conformation of the dinucleotides and the (6-4) PP formation mechanism, the key criteria to take into consideration is Φ since it is the first parameter that intervenes in the PP formation mechanism. Therefore, peak areas of PPs of **1** and **3**-**5** were measured at t ≤ 4 min and known Φ at 254 nm for TpT-derived (6-4) PP and CPD[18] were used to determine the quantum yields of the PPs of **3**-**4** (Table 2).[15c] Dinucleotides **3** and **4** were more prone to provide CPD than **1** (entries 3 and 4). In contrast, **3** and **5** were less efficient at providing (6-4) PP than **1** whereas $\Phi_{(6\text{-}4)}$ of dinucleotides **1** and **4** were almost similar (Table 2). At t ≤ 4 min, the comparative efficiency of (6-4) production from **1**, **3**-**5** determined by Φ followed the order of efficiency derived from comparison of relative chemical yields of the slope of the plot of relative chemical yields versus irradiation time.

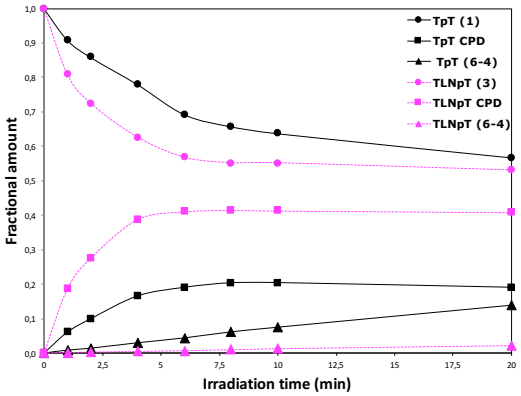


Fig. 3 Fractional amount of the $T_{LN}pT$ (**3**), TpT (**1**) and of their respective photoproducts as a function of irradiation time at 254 nm (monitored at 230 nm).

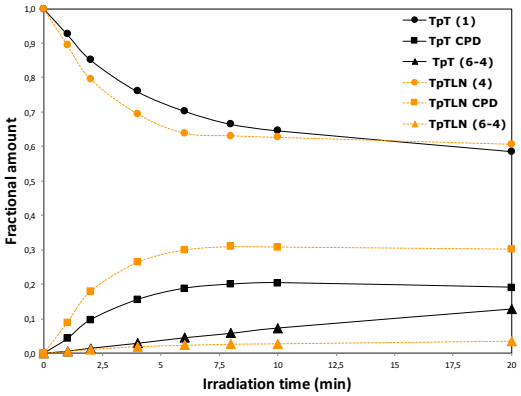


Fig. 4 Fractional amount of the TpT_{LN} (**4**), TpT (**1**) and of their respective photoproducts as a function of irradiation time at 254 nm (monitored at 230 nm).

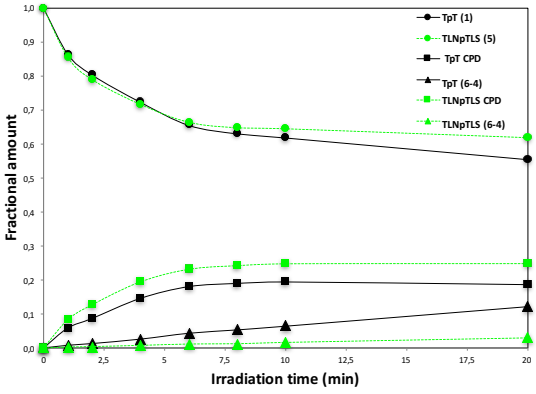


Fig. 5 Fractional amount of the $T_{LN}pT_{LS}$ (**5**), TpT (**1**) and of their respective photoproducts as a function of irradiation time at 254 nm (monitored at 230 nm).

This journal is © The Royal Society of Chemistry 20xx

Table 2 Quantum yields of PP formation, sugar conformational preferences (%), and CD-calculated stacking (%)[20] of dinucleotides 1-5

Entry	Compound	Φ (6-4).10^{-3}	Φ CPD.10^{-2}	Conformer blend				Stacking
				NN	NS	SN	SS	
1	TpT (**1**)[18]	1 ± 0.05	1.1 ± 0.05	11.5	16.5	29.5	42.5	24
2	$T_{LN}pT_{LN}$ (**2**)[15d]	0	6.8 ± 0.8	100	0	0	0	70
3	$T_{LN}pT$ (**3**)	0.20 ± 0.02	3.84 ± 0.26	39	61	0	0	43
4	TpT_{LN} (**4**)	0.80 ± 0.07	2.31 ± 0.15	37	0	63	0	26
5	$T_{LN}pT_{LS}$ (**5**)	0.33 ± 0.02	1.65 ± 0.11	0	100	0	0	22

This confirmed that, at t ≤ 4 min, no significant secondary photoreaction occurs since, in this case, the rate of photoproduct formation is related to the quantum yield (equation 4' ref 15c).

Overall, this result is in line with the notion of conformational rigidity that reduces (6-4) PP formation. Indeed, transforming the flexible sugar moiety into a rigid bicycle reduces the number of conformational degrees of freedom for the sugar phosphate backbone and therefore the conformational space available for the two proreactive thymine residues.

Combination of sugar conformers and their population in **1**-**5** are compiled Table 2. These populations encompass stacked (photoreactive) and unstacked (non photoreactive) species. Herein, the term "stacked" refers to bases sufficiently close in space and in a geometrical arrangement adequate to react after UV excitation unlike the "CD stacked species" that depends on the distance and dihedral angle between the two nucleobase moieties which are not necessarily those required for photoreaction. To avoid confusion, the term "photoreactive conformer" will be used in place of "stacked conformer" when appropriate.

CPD formation. Dinucleotide **2** ($T_{LN}pT_{LN}$) and **3** ($T_{LN}pT$) were calculated to be 100% NN and 61% NS (ie 39% NN), respectively (Table 2, entries 2 and 3). Quantum yield of CPD formation of **2** and **3** were determined to be $\Phi_{CPD} = 6.8 \times 10^{-2}$ and 3.8×10^{-2}, respectively (Table 2, entries 2 and 3). Considering that a portion of CPD issued from **3** had to be attributed to its NN population, this clearly indicated that the NS conformer is much less conducive to produce CPD than the NN conformer. This means that the thymine moieties in the NN conformer are more appropriately mutually oriented to give rise to CPD than in the NS conformer. This result was also corroborated by the low Φ_{CPD} value of the NS-dilocked compound $T_{LN}pT_{LS}$ (**5**, $\Phi_{CPD} = 1.6 \times 10^{-2}$, Table 2, entry 5; 100% NS, Table 2, entry 5), compared to that of **2**. In addition, stiffening the natural 3'-S conformer as in $T_{LN}pT_{LS}$ **5** did not result in an important variation of Φ_{CPD} compared to **3**, taking its NS conformer population (61%, Table 2, entry 3) into account. Results obtained for TpT_{LN} (**4**, $\Phi_{CPD} = 2.3 \times 10^{-2}$; 63% SN, Table 2, entry 4), taking into account the participation of its 37% NN population in CPD formation, indicated that in the SN conformer, the thymine moieties are not appropriately mutually-oriented to give rise to CPD. It has been previously reported[15d,g] that Φ_{CPD} is linearly correlated with the CD-derived intramolecular stacking level. Accordingly, **3** that was more efficient than **1** and less efficient than **2** at providing CPD, exhibited a stacking level of 43%, a value intermediate between that of **1** and **2** (24 and 70%, respectively) (Table 2, entries 1-3). This observation can be rationalized if the ground state of the photoreactive NN and NS conformers mainly participates in the CD signal of the population of stacked conformers.

These experimental results assert the conclusions drawn from theoretical results suggesting the NN and NS TpT conformers to be CPD-proreactive, and excluding the involvement of the SN conformer in CPD formation.[16a] Moreover, the NN and NS proreactive conformers were shown to exhibit the most intense calculated positive CD band absorption compared to the one of the stacked SN and SS conformers.[16a] This supports the correlation between Φ_{CPD} and CD-deduced stacking level hypothesis.

(6-4) PP formation. The (6-4) adduct formation is known to be prevented by freezing both sugar puckers of a thymine dinucleotide in the N conformation, as in $T_{LN}pT_{LN}$ (**2**).[15d] In contrast to **2**, locking only the 5'-N sugar conformation, as in $T_{LN}pT$ (**3**), allowed the production of (6-4) PP albeit in low yield ($\Phi_{(6-4)} = 0.2 \times 10^{-3}$, Table 2, entries 2 and 3). This result demonstrated that, among the two conformer populations of

3 (Table 2, entry 3), its NS population is able to afford (6-4) PP. The $\Phi_{(6\text{-}4)}$ of **3**, which is entirely due to its NS population (61%, Table 2, entry 3), is ca one fifth of **1** that contains only 16.5% of NS conformer ($\Phi_{(6\text{-}4)}$ = 1x10^{-3}, Table 2, entry 1). Therefore, the (6-4) PP formation capability of **1** must be mainly attributed to its SN and/or SS population of conformers, the NN conformer being non (6-4) photoreactive. Accordingly, the $\Phi_{(6\text{-}4)}$ of the NS-dilocked dinucleotide **5** ($T_{LN}pT_{LS}$) was calculated to be 0.33x10^{-3} (Table 2, entry 5), a value expected for a dinucleotide in an exclusive NS conformation considering the $\Phi_{(6\text{-}4)}$ of **3** and its NS population. Locking only the N sugar pucker at the 3'-end as in TpT_{LN} (**4**) also allowed the production of (6-4) PP ($\Phi_{(6\text{-}4)}$ = 0.8x10^{-3}, Table 2, entry 4). Moreover, since the population of the SN conformer of **4** and NS conformer of $T_{LN}pT$ **3** are similar (Table 2, entries 3 and 4), the SN conformer can be estimated to be ca four-fold more prone to give rise to (6-4) PP than the NS conformer **3** (Table 2, entries 3 and 4). Such finding supports our hypothesis that the SN conformer fraction of **1** is important for (6-4) adduct production. Formation of (6-4) PP and level of stacked species observable by CD are not positively linearly correlated, as previously reported (Table 2).[15g] This is likely because the ground state of the proreactive SN conformer does not participate strongly to the CD signal due to its geometry.

Our results constitute the first experimental proof of the involvement of the SN conformer in providing (6-4) PP as previously suggested by theoretical studies.[16a] Moreover, this conformer has been shown to exhibit a calculated positive CD band absorption of lower amplitude than the one of the NN stacked conformer[16a] in line with our proposed rationalization of the absence of correlation between $\Phi_{(6\text{-}4)}$ and CD-deduced stacking level. Our study also experimentally demonstrates that the NS conformer is capable of (6-4) PP production as well, even though moderately with respect to the SN conformer. Such finding is of particular interest to the light of theoretical studies hypothesizing this conformer to be unprone to oxetane formation.[16a]

Conclusions

Undoubtedly, independently on the type of sugar pucker, the CPD formation-efficiency of **3**-**5** consistently remains higher than that of the (6-4) PP. In addition, the impact of puckering variation is largely reflected on the CPD formation leading to experimentally identify the NS conformer as a new pro-CPD species. Even if the (6-4) PP formation ability of **3**-**5** is much lower than that of the CPD, the biological relevance of (6-4) PPs is such that any modulation in their production has to be comprehensively explored.

Previous experimental photochemical studies had shown that the S conformation of at least one sugar residue in TT site was crucial for (6-4) PP formation.[15d,f,g] Herein, we experimentally demonstrate that this S-puckered sugar can be located either at the 5'- or 3'-end of the TT site. Moreover, we establish that the 5'-end S conformer location is the most efficient at providing (6-4) PP. In addition, specificity (with regard to CPD production) is obtained with the 5'-end S puckered conformer. It now remains to experimentally establish if two consecutive S conformers would be prone to trigger (6-4) adduct formation.

Beyond increasing our intimate knowledge on the formation mechanism of UV-induced TpT PPs and laying the foundation for future biophysical, photochemical and theoretical work as $T_{LN}pT_{LN}$ (**2**) inspired,[15e,f,16a] our studies allow dissecting the correlation between sugar phosphate backbone structural parameters and geometric stacking patterns for the dipyrimidine dinucleotide motif in solution. Indeed, three out of the four combinations of sugar puckers in the dithymine dinucleotide single strand can now be related to pro photoreactive stacking patterns compatible with (6-4) and/or CPD production.

Interestingly, various thymine thymine dinucleotide stacking geometries have been identified in single-stranded regions of nucleic acid crystal structures.[34,35] Our results are likely to be useful to identify staked state candidates as proreactive species based on their sugar pucker combination.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank CNRS and Conseil Régional Champagne Ardenne for a doctoral fellowship to J. J. (completed by Fondation pour la Recherche Médicale (FDT20130928264)) and for their financial support to the PlAneT CPER project co-funded by Ministère de l'Enseignement Supérieur et de la Recherche and EU-program FEDER.

References

1 (a) J. Cadet, S. Courdavault, J.-L. Ravanat and T. Douki, *Pure Appl. Chem*., 2005, **77**, 947-961. (b) J. Cadet, S. Mouret, J.-L. Ravanat and T. Douki, *Photochem. Photobiol*., 2012, **88**, 1048-1065.

2 (a) G. P. Pfeifer, *Genome Instab. Dis.*, 2020, **1**, 99-113. (b) L. H. F. Mullenders, *Photochem. Photobiol. Sci.*, 2018, **17**, 1842-1852. (c) J. Cadet and T. Douki, *Photochem. Photobiol. Sci.* 2018, **17**, 1816-1841. (d) N. C. Delic, J. G. Lyons, N. D. Girolamo and G. M. Halliday, *Photochem. Photobiol.*, 2017, **93**, 920-929. (e) H. Kaddurah, T. L. Braunberger, G. Vellaichamy, A. F. Nahhas, H. W. Lim and I. H. Hamzavi, *Curr. Geri. Rep*. **2018**, *7*, 228-237.

3 (a) C. L. Ballaré, M. M. Caldwell, S. D. Flint, S. A. Robinson and J. F. Bornman, *Photochem. Photobiol. Sci.*, 2011, **10**, 226-241. (b) N. Tuteja, P. Ahmad, B. B. Panda and R. Tuteja, *Mut. Res.*, 2009, **681**, 134-149.

4 W. J. Zhang and L. O. Björnrn, *Fitoterapia*, 2009, **80**, 207-218.

5 (a) T. Douki and J. Cadet, *Biochemistry*, 2001, **40**, 2495-2501. (b) S. Mouret, C. Baudouin, M. Charveron, A. Favier, J. Cadet and T. Douki, *Proc. Natl. Acad. USA*, 2006, **103**, 13765-13770.

6 (a) J. Hu, S. Adar, C. P. Selby, J. D. Lieb and A. Sancar, *Genes Dev.*, 2015, **29**, 948-960 and references cited. (b) S. Courdavault, C. Baudouin, M. Charveron, B. Canguilhem, A. Favier, J. Cadet and T. Douki, *DNA Repair*, 2005, **4**, 836-844.

7 (a) S. Nakajima, L. Lan, S.-I. Kanno, M. Takao, K. Yamamoto, A. P. M. Eker and A. Yasui, *J. Biol. Chem*., 2004, **279**, 46674-

 This journal is © The Royal Society of Chemistry 20xx

46677. (b) H. L. Lo, S. Nakajima, L. Ma, B. Walter, A. Yasui, D. W. Ethell and L. B. Owen, *BMC Cancer*, 2005, **5**, 135. (c) E. Otoshi, T. Yagi, T. Mori, T. Matsunaga, O. Nikaido, S.-T. Kim, K. Hitomi, M. Ikenaga and T. Todo, *Cancer Res*. 2000, **60**, 1729-1735. (d) J.-I. Akagia, K. Hashimotoc, K. Suzukia, M. Yokoi, N. de Winde, S. Iwai, H. Ohmoria, M. Moriyac and F. Hanaokaa, *DNA Repair*, 2020, **87**, 102771 and references cited.

8 K.-F. Hung, J. M. Sidorova, P. Nghiem and M. Kawasumi, *Proc. Natl. Acad. USA*, 2020, **117**, 12806-12816.

9 For recent studies see: (a) A. Giussani and G. A. Worth, *J. Phys. Chem. Lett*., 2020, **11**, 4984-4989. (b) J. Gontcharov, L. Liu, B. M. Pilles, T. Carell, W. J. Schreier and W. Zinth, *Chem. Eur. J*., 2019, **25**, 15164-15172. (c) J.-H. Li, T. J. Zuehlsdorff, M. C. Payne and N. D. M. Hine, *J. Phys. Chem. C*, 2018, **122**, 11633–11640. (d) A. Giussani, I. Conti, A. Nenov and M. Garavelli, *Faraday Discuss*., 2018, **207**, 375-387.

10 W. J. Schreier, T. E. Schrader, F. O. Koller, P. Gilch, C. E. Crespo-Hernández, V. N. Swaminathan, T. Carell, W. Zinth and B. Kohler, *Science*, 2007, **315**, 625-629.

11 P. Clivio, J.-L. Fourrey, J. Gasche and A. Favre, *J. Am. Chem. Soc*., 1991, **113**, 5481-5483.

12 L. A. Ortiz-Rodríguez, C. Reichardt, S. J. Hoehn, S. Jockusch and C. E. Crespo-Hernández, *Nature Commun*., 2020, **11**, 3599.

13 (a) V. Lyamichev, M. D. Frank-Kamenetskii, and V. N. Soyfer, *Nature*, 1990, **344**, 568-570. (b) L. M. Kundu, U. Linne, M. Marahiel and T. Carell, *Chem. Eur. J.*, 2004, **10**, 5697-5705. (c) T. Douki, *J. Photochem. Photobiol. B: Biol.*, 2006, **82**, 45-52. (d) M. McCullagh, M. Hariharan, F. D. Lewis, D. Markovitsi, T. Douki and G. C. Schatz, *J. Phys. Chem. B*, 2010, **114**, 5215-5221. (e) Z. Pan, M. McCullagh, G. C. Schatz and F. D. Lewis, *J. Phys. Chem. Lett*., 2011, **2**, 1432-1438. (f) M. Hariharan, K. Siegmund, C. Saurel, M. McCullagh, G. C. Schatz and F. D. Lewis, *Photochem. Photobiol. Sci*., 2014, **13**, 266-271.

14 (a) P. Mao, J. J. Wyrick, S. A. Roberts and M. J. Smerdon, *Photochem. Photobiol*. 2017, **93**, 216-228 and references cited. (b) J. Hu, O. Adebali, S. Adar and A. Sancar, *Proc. Natl. Acad. USA*, 2017, **114**, 6758-6763. (c) Y. Wang, M. L. Gross, and J.-S. Taylor, *Biochemistry*, 2001, **40**, 11785-11793.

15 (a) T. Ostrowski, J.-C. Maurizot, M.-T. Adeline, J.-L. Fourrey and P. Clivio, *J. Org. Chem.*, 2003, **68**, 6502-6510. (b) G. P. H. Santini, C. Pakleza, P. Auffinger, C. Moriou, A. Favre, P. Clivio and J. A. H. Cognet, *J. Phys. Chem. B*, 2007, **111**, 9400-9409. (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. (d) C. Desnous, B. R. Babu, C. Moriou, J. U. Ortiz Mayo, A. Favre, J. Wengel and P. Clivio, *J. Am. Chem. Soc*., 2008, **130**, 30-31. (e) W. J. Schreier, J. Kubon, N. Regner, K. Haiser, T. E. Schrader, W. Zinth, P. Clivio and P. Gilch, *J. Am. Chem. Soc*., 2009, **131**, 5038-5039. (f) M. Hariharan, M. McCullagh, G. C. Schatz and F. D. Lewis, *J. Am. Chem. Soc*., 2010, **132**, 12856-12858; corrected *J. Am. Chem. Soc*., 2010, **132**, 15831. (g) C. Moriou, C. Denhez, O. Plashkevych, S. Coantic-Castex, J. Chattopadhyaya, D. Guillaume and P. Clivio, *J. Org. Chem.*, 2015, **80**, 615-619. (h) C. Moriou, A. D. Da Silva, M. J. Vianelli Prado, C. Denhez, O. Plashkevych, J. Chattopadhyaya, D. Guillaume and P. Clivio, *J. Org. Chem*., 2018, **83**, 2473-2478.

16 (a) R. Improta, *J. Phys. Chem. B*, 2012, **116**, 14261-14274. (b) I. Conti,; L. Martínez-Fernández, L. Esposito, S. Hofinger, A. Nenov, M. Garavelli and R. Improta, *Chem. Eur. J*., 2017, **23**, 15177-15188. (c) L. Martínez-Fernández and R. Improta, *Photochem. Photobiol. Sci*., 2018, **17**, 586-591.

17 T. Douki, M. Court and J. Cadet, *J. Photochem. Photobiol. B: Biol*., 2000, **54**, 145-154.

18 H. E. Johns, M. L. Pearson, J. C. LeBlanc and C. W. Helleiner, *J. Mol. Biol*., 1964, **9**, 503-524.

19 C. Nganou, S. D. Kennedy and D. W. McCamant, *J. Phys. Chem. B*, 2016, **120**, 1250-1258.

20 J. Jakhlal, S. Coantic-Castex, C. Denhez, C. Pertermann, A. Martinez, D. Harakat, D. Guillaume and P. Clivio, *Chem. Commun.*, 2015, **51**, 12381-12383.

21 (a) A. A. Koshkin, V. K. Rajwanshi and J. Wengel, *Tet. Lett.*, 1998, **39**, 4381-4384. (b) A. A. Koshkin, S. K. Singh, P. Nielsen, V. K. Rajwanshi, R. Kumar, M. Meldgaard, C. E. Olsen and J. Wengel, *Tetrahedron*, 1998, **54**, 3607-3630. (c) A. A. Koshkin, J. Fensholdt, H. M. Pfundheller and C. Lomholt, *J. Org. Chem*., 2001, **66**, 8504-8512.

22 (a) S. Obika, K. Morio, Y. Hari and T. Imanishi, *Chem. Commun*., 1999, 2423-2424. (b) S. Obika, K. Morio, D. Nanbu, Y. Hari, H. Itoh and T. Imanishi, *Tetrahedron*, 2002, **58**, 3039-3049.

23 S. Obika, T. Uneda, T. Sugimoto, D. Nanbu, T. Minami, T. Doi and T. Imanishi, *Bioorg. Med. Chem*., 2001, **9**, 1001-1011.

24 P. Savy, R. Benhida, J.-L. Fourrey, R. Maurisse and J.-S. Sun, *Bioorg. Med. Chem. Lett*., 2000, **10**, 2287-2289.

25 (a) A. Kumar and R. T. Walker, *Tetrahedron*, 1990, **46**, 3101-3110. (b) R. Chung and K. A. Anderson, *Tetrahedron Lett*., 2006, **47**, 8361-8363.

26 J. D. More and N. S. Finney, *Org. Lett.*, 2002, **4**, 3001–3003.

27 P. M. S. Hendrickx and J. C. Martins, *Chem. Cent. J*., 2008, **2**, 20.

28 O. Plashkevych, S. Chatterjee, D. Honcharenko, W. Pathmasiri and J. Chattopadhyaya, *J. Org. Chem*., 2007, **72**, 4716-4726.

29 K. Morita, M. Takagi, C. Hasegawa, M. Kaneko, S. Tsutsumi, J. Sone, T. Ishikawa, T. Imanishi and M. Koizumi, *Bioorg. Med. Chem*., 2003, **11**, 2211-2226.

30 (a) C. B. Nielsen, S. K. Singh, J. Wengel and J. P. Jaconsen, *J. Biomol. Struct. Dyn*., 1999, **17**, 175-191. (b) M. Petersen, K. Bondensgaard, J. Wengel and J. P. Jacobsen, *J. Am. Chem. Soc*., 2002, **124**, 5974-5982. (c) K. Bondensgaard, M. Petersen, S. K. Singh, V. K. Rajwanshi, R. Kumar, J. Wengel and J. P. Jacobsen, *Chem. Eur. J.*, 2000, **6**, 2687-2695. (d) M. Petersen, C. B. Nielsen, K. E. Nielsen, G. A. Jensen, K. Bondensgaard, S. K. Singh, V. K. Rajwanshi, A. A. Koshkin, B. M. Dahl, J. Wengel and J. P. Jacobsen, *J. Mol. Rec.*, 2000, **13**, 44-53. (e) A. Ivanova and N. Rösch, *J. Phys. Chem. A*, 2007, **111**, 9307-9319. (f) G. Suresh and U. D. Priyakumar, *J. Phys. Chem. B*, 2014, **118**, 5853-5863.

31 M. Egli, G. Minasov, M. Teplova, R. Kumar and J. Wengel, *Chem. Commun*., 2001, 651-652.

32 C. Altona, *Recl. Trav. Chim. Pays-Bas*, 1982, **101**, 413-433.

33 Y. Wang, J.-S. Taylor and M. L. Gross, *Chem. Res. Toxicol.*, 1999, **12**, 1077-1082.

34 A. Sedova and N. K. Banavali, *Biochemistry*, 2017, **56**, 1426-1443.

35 J. Černý, P. Božíková, J. Svoboda and B. Schneider, *Nucleic Acids Res*., 2020, **48**, 6367-6381.