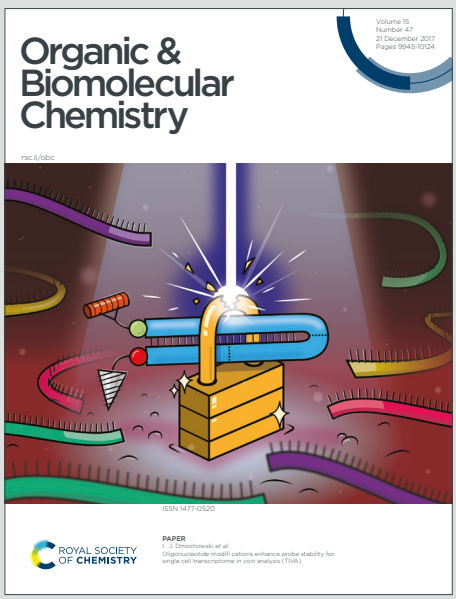
Organic & Biomolecular Chemistry

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: M. TAVERNA PORRO, C. Saint-Pierre, D. Gasparutto and J. Ravanat, *Org. Biomol. Chem.*, 2020, DOI: 10.1039/C9OB01021A.



ROYAL SOCIETY OF CHEMISTRY

rsc.li/obc

View Article Online
DOI: 10.1039/C9OB01021A

ARTICLE

Solid-phase synthesis of branched oligonucleotides containing the biologically relevant dCyd341 interstrand crosslink DNA lesion†.

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

DOI: 10.1039/x0xx00000x

Marisa L. Taverna Porro[a,b], Christine Saint-Pierre[a], Didier Gasparutto[a] and Jean-Luc Ravanat[a]*

Branched oligonucleotides containing a biological relevant DNA lesion, dCyd341 that involves an interstrand crosslink between a cytosine base on one strand and a ribose moiety of the opposite strand, was prepared in a single automated solid-phase synthesis. For this, we first prepared the phosphoramidite analogue of dCyd341 bearing an orthogonal levulinyl protecting group. Then, following synthesis of the first DNA strand containing dCyd341, levulinic group was removed and synthesis was then continued from the free base hydroxyl group at the branching point, using traditional phosphoramidites. The synthesized oligonucleotides were fully characterized by MALDI-TOF/MS, were enzymatically digested and the presence of the lesion was confirmed by LC/MS-MS and sequence was finally controlled upon exonucleases digestion followed by MALDI-TOF/MS analysis. The developed strategy was successfully employed for the preparation of several short linear and branched oligonucleotides containing the aforementioned lesion.

Introduction

Oxidative stress is known to be implicated in several biological disorders and associated with diverse pathologies.[1-3] In particular, reactive oxygen species are known to induce chemical modifications to biomolecules. DNA is one of the main cellular targets of oxidative stress and the produced DNA lesions can contribute to the initiation and/or propagation processes that can lead to carcinogenesis through mutagenesis.[4] To reduce the deleterious effects of these modifications, cells have developed efficient DNA repair systems able to excise these lesions and restore DNA integrity.

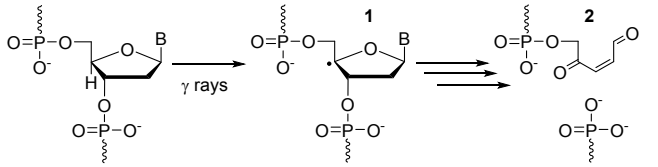
Complex or clustered DNA lesions, including tandem lesions, double strand breaks and interstrand crosslinks (ICLs)[5] involve several chemical modifications of the DNA structure in one helix turn.[2,6] For ICLs the covalent tethering of both duplex strands prevents their separation; blocking essential biological processes such as DNA replication and transcription.[7] Thus, ICLs represent a challenge for the DNA repair machinery. Their repair involves nucleotide excision repair (NER), homologous recombination (HR) and non-homologous end joining (NHEJ) pathways, but a lot is still unclear.[7] If left unrepaired, they can provoke chromosomal breakages, rearrangements, genomic instability and aging in tissue and organs.[8] Such a toxic effect is the principle of action of many antitumoral drugs currently used in cancer therapy.[9]

The ICLs lesions can be exogenously formed by action of chemical agents such as bifunctional alkylating agents including platinum compounds and psolaren.[10] Recent results[11] have indicated that they could also be generated as a consequence of a single ionization process, and therefore by endogenous oxidative stress, ionizing radiation and radiomimetic drugs. This concerns lesions generated through the transient formation of reactive intermediates (aldehydes or radicals) on one DNA strand, that may further react with surrounding DNA bases.[12]

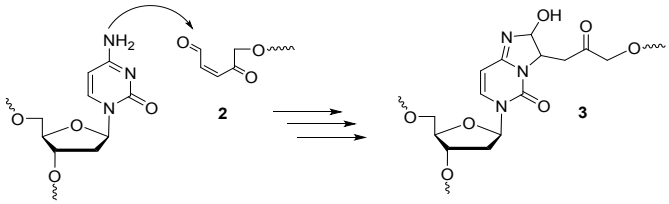
Regarding this last type of damage, a lesion named dCyd341, identified by our laboratory, was found to be generated in significant amounts in cellular DNA even by endogenous oxidative stress,[13] whereas its repair rate (half-life of about 10h in culture cells) was rather slow compared to classical oxidative DNA lesions. It was clearly demonstrated that dCyd341 is a complex DNA lesion generated by a single ionization process (scheme 1).[13] Indeed, the initial event leading to formation of the lesion involves hydrogen abstraction at the C4´ position of the 2-deoxyribose moiety of DNA with the concomitant formation of the corresponding C4' radical. Decomposition of that radical (**1**, Scheme 1) gives rise to the formation of a modified sugar residue bearing a reactive aldehyde (**2**, Scheme 1). This transiently generated aldehyde is able to add onto a cytosine base (and also adenine[14]) located onto the opposite strand thus generating through successive steps to the four possible diastereomers of dCyd341 (Fig S1). Such a mechanism generates a strand break and an interstrand crosslink (Schemes 1 and 2).

[a] Univ. Grenoble Alpes, CEA, CNRS IRIG/SyMMES, F-38054 Grenoble Cedex 9.
[b] Present Address: [Marisa L. Taverna Porro], IQUIMEFA, FFyB, UBA, Buenos Aires, Argentina
[c] * To whom correspondence should be addressed. Tel +33 4 38 78 47 97 ; Email: jravanat@cea.fr

†Electronic supplementary information (ESI) available: NMR spectra for all the new compounds, HPLC-MS/MS profiles of deprotected compound **6** and Maldi-TOF/MS analysis of synthetized oligonucleotides and their digestion products obtained by exonucleases hydrolysis.



Scheme 1. Formation of reactive aldehyde **2** by C4´-hydrogen abstraction reaction.



Scheme 2. Formation of dCyd341 (**3**) initiated by nucleophilic addition of the amino group of cytosine onto the reactive aldehyde **2**.

Therefore, it is of particular interest to determine i) the biological relevance of this lesion, ii) its kinetics and efficacy of elimination by DNA repair processes and iii) its mutagenic potential. It is well known that the extensive study of ICLs is hampered by the difficulty in the preparation of suitable substrates that are required to perform biochemical studies. One of the biggest challenges in the characterization of ICLs at the molecular level is the preparation of short oligonucleotide (ON) duplexes containing a single cross-link at a specified and well-defined location in the DNA sequence.

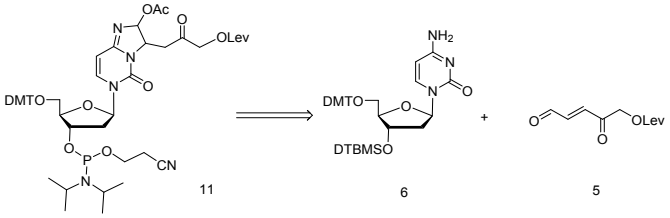
Nevertheless, several strategies for their preparation have already been investigated. For example, direct treatment of DNA with bifunctional agents has been reported.[15,16] This strategy often gives rise to a complex spectrum of products and the desired modification could only be isolated in small yields. Another methodology involves the construction of short DNA sequences bearing a reactive residue (or an activatable one) towards the complementary strand. Even if such a strategy was employed successfully in the synthesis of several short ICLs duplexes, it presents a general disadvantage of lack of specificity and instability (to see some examples[17-21]). Another strategy consists in using an orthogonal protecting group. This approach involves the synthesis of DNA oligomers on a solid-phase automated DNA synthetizer, incorporating a modified nucleoside phosphoramidite with an orthogonal protecting group as the branching point. Once the first DNA strand is synthesized, the orthogonal protecting group can be selectively removed leaving the protected DNA strand attached on the solid support. The branched sequence can then be completed from the unprotected branching point, using the solid-phase automated DNA synthetizer. This methodology has been successfully applied for site specific synthesis of RNA[22] and DNA.[23]

In this work, we report the first synthesis of branched oligonucleotides containing the biological relevant dCyd341 lesion. For this, we have successfully prepared the dCyd341 phosphoramidite analogue bearing an orthogonal protecting group at the branching point. It was then used for the preparation of short linear and branched oligonucleotides containing the aforementioned DNA lesion.

Results and discussion

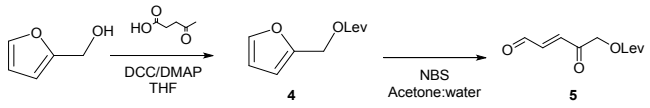
Synthesis of dCyd341 phosphoramidite

The targeted phosphoramidite synthon (**11**, Scheme 3) for this strategy is shown in Scheme 3. The levulinyl group was chosen as the orthogonal protecting group since, as previously reported,[24] it can be easily removed using hydrazine without affecting the other protective functions of the DNA oligomer. The base modification can be achieved, as already described by our group[13], by reaction of the protected nucleoside with the levulinated conjugated aldehyde. Therefore, the first step of this synthetic strategy was the preparation of the corresponding aldehyde (**5**, Scheme 3).



Scheme 3. Synthetic strategy of phosphoramidite synthon **11**.

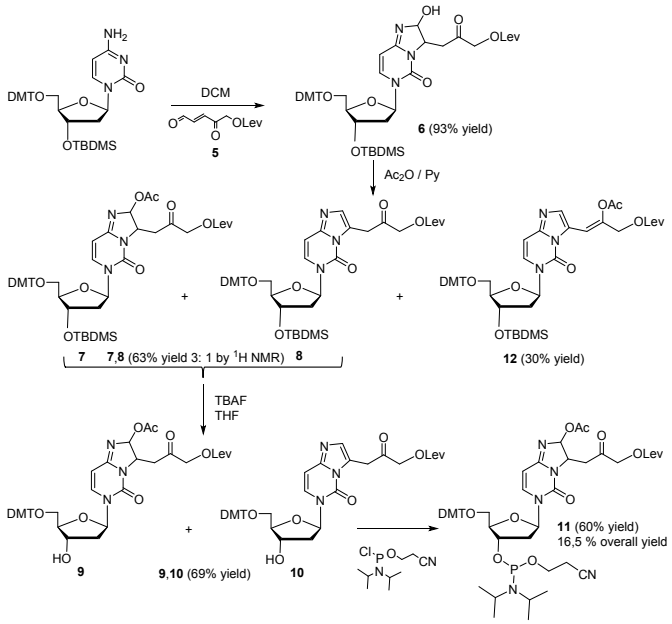
For this, furfuryl alcohol was acylated using levulinic acid and DCC to give furfurylllevulinate (**4**, Scheme 4) in quantitative yields. Oxidation of **4** was then investigated. Then, the protected furfuryl derivative was oxidized using different oxidizing reagents, namely ozone, 3-chloroperoxybenzoic acid (MCPBA) and N-bromosuccinimide (NBS). When MCPBA was used only 8% of the furan moiety was converted into the aldehyde after 2 days of reaction. Ozone-mediated oxidation of **4** led to total conversion of the starting material, but only small amounts of the desired aldehyde were recovered. NBS oxidation gave the best results, generating the trans 2,5-dioxo-pent-3-enyl levulinate (**5**, scheme 4) with 75% yield in 3 hours.



Scheme 4. Chemical synthesis of 2,5-dioxo-pent-3-enyl levulinate **5**.

The reaction mixture of **5** was then incubated in the presence of the protected deoxycytidine to give the adduct **6** in high yields (90 %, Scheme 5). The reaction is supposed to occur by initial attack of the exocyclic amine followed by 1,4-addition of the adjacent endocyclic nitrogen atom to the double bond, to generate an oxadiazabicyclo(3.3.0)octaimine adduct **6**. Even though mass spectrometry analysis of the purified adduct exhibit the expected molecular weight (m/z = 856 Da, positive mode) this compound could not be unambiguously characterized by NMR, even using 2D COSY, HSQC and HMBC NMR spectra (Fig S2, S3, S4, S5, S6) due to the presence of several diastereomers[5] (chemical structures of all

possible diastereomers are presented in Fig. S1). However, high resolution mass spectrometry of compound **6** (Fig. S7) is in agreement with the proposed structure. For further confirmation, the oxidized furfuryl levulinate derivative **5** was incubated in the presence of 2-deoxycytidine. The obtained products were then treated under alkaline conditions to remove the levulinyl group and analyzed by HPLC-MS/MS. The synthesized products, which appear as four modified nucleosides, were found to exhibit the same HPLC retention times and mass spectroscopic features as the four isomers of dCyd341, whose structures has been previously characterized in our laboratory [5,13] (Fig. S8).



Scheme 5. Synthesis of dCyd341 phosphoramidite **11**.

The modified nucleoside **6** was then acetylated using acetic anhydride in pyridine and the reaction mixture was then purified by silica gel chromatography. Two major fractions were collected and analyzed by mass spectrometry. ESI$^+$ MS analysis of the first fraction revealed a major compound with m/z=880 Da (**12**, Scheme 5), whereas the second fraction contained a mixture of two products with molecular weights of 837 (**8,** Scheme 5) and 897 (**7**, Scheme 5), the last being the molecular weight of the desired acetylated compound. Further insight into the structures of these compounds was obtained from tandem mass spectometry and ^{1}H-NMR. The MS/MS fragmentation patterns (Fig S9) presented the predominant loss of the protected sugar, confirming that the difference in mass could be attributed to the base moieties (m/z = 348, 306 and 368 for **12**, **8** and **7,** respectively). MS^2 fragmentation of **12** exhibited a major secondary product ion at m/z = 306 and the same MS^3 fragmentation pattern to the **8** BH^{2+} fragment, suggesting that the differences between these two compounds was an acetyl group. The Δm/z between **12** and **7** was 18 Da, indicating that **12** could be formed by a dehydration reaction of compound **7** (Scheme 5). Efforts were then made to separate small amounts of each compound presented in the second fraction. Although these compounds could not be fully separated, small amounts of compound **8** was recovered and characterized by ^{1}H-NMR. The aromatic region of the ^{1}H-NMR spectra recorded for compounds **12** and **8** was characterized by two singlets at around 7.63 ppm (B, Fig. 1) and one singlet at 7.22 ppm (A, Fig. 1) respectively, and these signals were missing for compound **7**. The ^{1}H NMR spectrum of **12** and **7** also showed the presence of the CH_3 signals of the acetyl group at around 2 ppm (not shown), whereas this signal was absent for compound **8**.

Taking all of the analysis into account, our results suggest that **7** is the desired acetylated adduct, **8** being its dehydrated product formed by the elimination of the acyloxy group elicited by the acetylation reaction. Its formation may be rationalized as an enolization followed by an acetyl migration and subsequent dehydration. The driven force for this last reaction could be the greater stability derived from the extended conjugation of the generated product. Other acetylation conditions as well as other acylating reagents were studied in order to avoid the formation of these by-products. Whatever the conditions used, the dehydration product **8** was always generated.

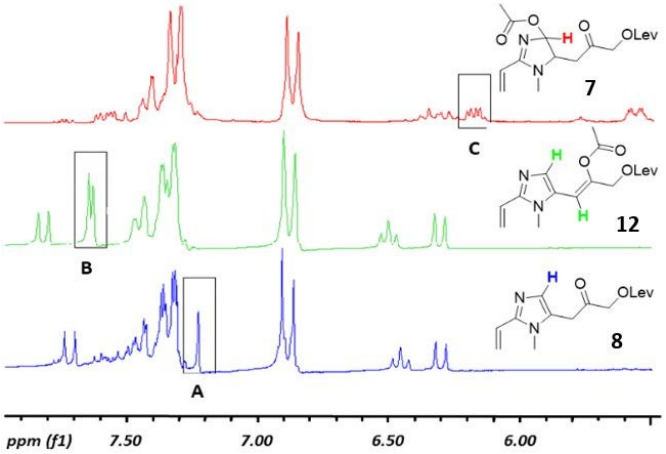


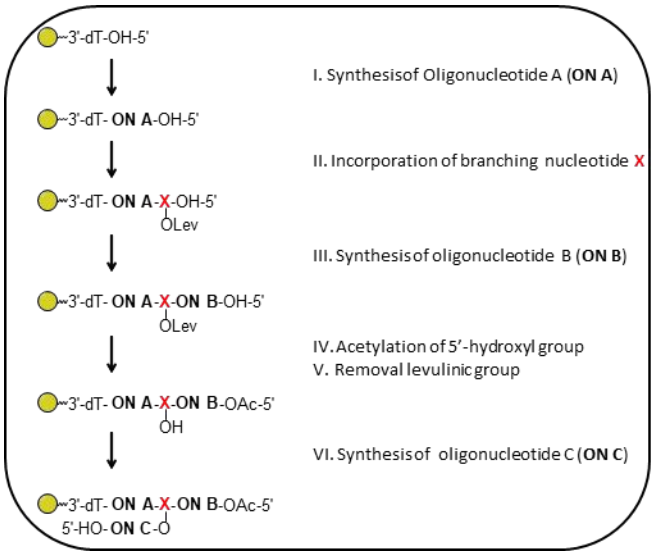
Fig. 1. Partial ^{1}H-NMR sepectra (5-8 ppm) of acetylation products **7** (red), **8** (blue) and **12** (green). Interesting protons are highlighed (see discussion).

The synthesis of the phoshoramidite was continued with the mixture of acetylated compounds **7** and dehydration product **8**. For this, the tert-butyldimethylsilyl (TBDMS) group was selectively removed by treating the mixture of **7** and **8** with tetrabutylammonium fluoride (TBAF) in THF, yielding a mixture of free 3´-hydroxy derivatives **9** and **10**. Finally, the resulting mixture was phosphitylated using excess of 2-cyanoethyl N,N-diisopropylchlorophosphoramidite. Fortunately, for compound 10, the oxygen atom connected as an enol to the aromatic base was also reactive towards the phosphitylating reagent, affording two chromatographically separable phosphoramidite derivatives, desired compound **11** (60 %, Scheme 5) and biphosphitylated compound **13** (Chemical structure presented in Fig. S10).

Oligonucleotide synthesis

The strategy for the solid-phase synthesis of the branched oligonucleotide is shown in Scheme 6. The first step involves the traditional solid-phase synthesis of oligonucleotides with the

incorporation of the modified phosphoramidite bearing the branching point. Once the linear strand is completed, an acetylation step is performed in order to block further elongation at the 5'-end of the oligomer. The protecting group of the branching point is then selectively removed and the synthesis of the branched oligonucleotide is continued.

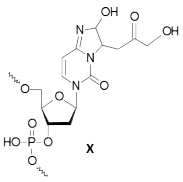


Scheme 6. Strategy for the synthesis of branched oligonucleotides bearing **dCyd341**.

First, in order to determine the ability to incorporate the modified phosphoramidite into oligonucleotides, regular non-branched oligomers of several lengths were prepared as listed in Table 1 (**14**, **15**, **17** and **18**). The modified oligonucleotides were synthesized using standard fast deprotection phosphoramidite chemistry with cycles including detritylation, coupling, capping and oxidation. The syntheses were initiated by constructing the linear strand using standard amidite concentration and coupling times. However, when the branching monomer phosphoramidite **11** was incorporated, a coupling time of 15 min was applied. Under these conditions, the coupling efficiency was almost quantitative. The oligonucleotides were then cleaved from the solid support and deprotected by treatment with concentrated ammonia for 5 h at room temperature (approximately yields 60 % for all oligonucleotides synthesized). The released oligonucleotides were purified by HPLC to give the deprotected oligonucleotides. These oligonucleotides were analyzed by MALDI-TOF/MS and the observed molecular weights supported the incorporation of the modified phosphoramidite onto the oligonucleotides (Fig S11 and Table 1).

In order to fully characterize the modified oligonucleotides, the purified oligomers were enzymatically digested and analyzed by HPLC/MS-MS. In all of the digested samples, for both linear and branched oligomers, a compound with a molecular mass corresponding to dCyd341 and at the expected retention time was observed (Fig. S12). No appreciable amount of the dehydration product was detected. For preparing the branched oligonucleotides, a separate end-capping cycle (consisting of detritylation and acetylation) was used to block further elongation at the 5'-end. Thereafter, the support was treated with a 0.5 M solution of hydrazine hydrate in pyridine:acetic acid for 15 min to remove the levulinyl protecting group. Synthesis was then continued in the normal fashion from the free base hydroxyl group at the branching point, using traditional phosphoramidites. The molecular weight of the ONs was determined by MALDI-TOF/MS (Fig. S11 and Table 1). As expected, both the non-branched **15** and branched oligonucleotides **16** of the same length share the same molecular weight.

Table 1. Oligonucleotides sequences and masses determined by MALDI-TOF/MS and compared to theoretical calculations.



ON # Sequence (5'-3')	MALDI-TOF/MS Cald. mass (Da)	Exp. mass (Da)
14 GTXAC	1577.1	1575.8
15 CTGTXAC	2170.5	2169.0
16 $^{5'}$CT / $^{5'}$GTXAC$_{3'}$	2170.5	2169.7
17 GTCAGGXAACGTC	3760.5	3759.6
18 GCGTTAGTCAGGXACTTCAG	6272.0	6270.8

Then, to verify the successful synthesis of the branched oligomers two different analyses were performed. For practical purposes, the analyses were carried out with the shortest sequences (**15** and **16**, Table 1). First, examination of the HPLC profiles of the purified branched and non-branched oligonucleotides reveals different chromatographic behaviors, both oligonucleotides having different retention times (A and B, Fig. 2). The co-injection of a mixture of them confirmed this assumption as shown in panel C (Fig. 2).

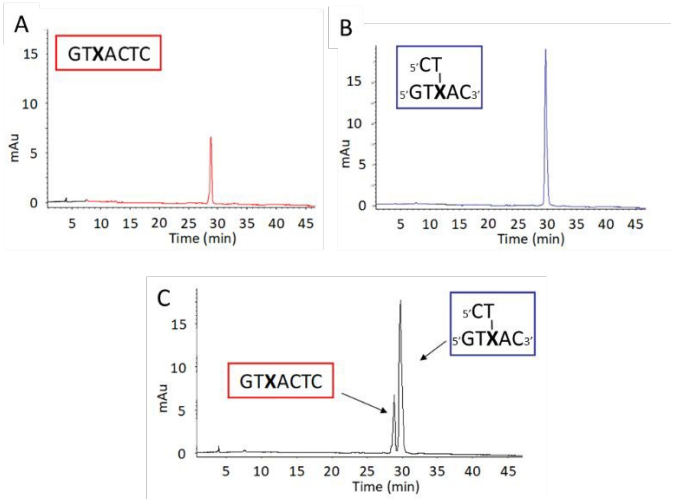


Fig. 2. HPLC chromatographic profiles of: A. **15**, B. **16** and C. Co-injection of **15** and **16**.

Finally, controlled phosphodiesterase digestions followed by MALDI-TOF/MS analysis was performed. Oligonucleotides were partially digested with 3' → 5' and 5' → 3' acting exonucleases in a MALDI-TOF/MS compatible buffer. The partial digests were analyzed by MALDI-TOF mass spectrometry and the sequence was inferred from the mass differences between adjacent peaks (Table 2).

Surprisingly, digestion of branched oligonucleotides was faster than that of the corresponding non-branched ones. The MALDI-TOF/MS profiles for the 3' → 5' digestion are, as expected, similar for both oligonucleotides (Table 2B and 2D, Fig. S13), showing both the sequential loss of dC and dA. For 5' → 3' digestion, the profiles are quite different (Table 2A and 2C, Fig. S14). For the non-branched oligonucleotide (Table 2A), the enzymatic hydrolysis proceeds up to one nucleotide before the modification, observing peaks at m/z = 1880, 1576 and 1245 corresponding to the sequential loss of dC, dT and dG, respectively. For the branched oligonucleotide (Table 2C) only two nucleotides were lost before the digestion dramatically stopped (m/z 1880, 1840 and 1550, Table 2). The peak at m/z = 1880 corresponds to the loss of one dG, only possible for the branched sequence whereas the peak at m/z = 1840 correspond to the loss of dC. The fragment observed at 1551, corresponding to the loss of both dG and dC is also only possible in a branched sequence.

Altogether, the obtained results clearly demonstrate that after deprotection of the orthogonal levulinyl group and further elongation, we can prepare the biologically relevant branched oligonucleotides containing the dCyd341 lesion.

Table 2. Application of exonuclease digestion together with MALDI-TOF/MS analysis for the characterization of branched (**16**) and non-branched oligonucleotides (**15**). MW observed for the: A. 5'-3' digestion of **15**, B.3'-5' digestion of **15**, C.5'-3' digestion of **16** and D. 3'-5' digestion of **16**.

A

5'-CTGTXAC-3'

5'-exonuclease	MALDI-TOF/MS Cald. mass	Exp. mass
5'-CTGTXAC-3'	2169,6	2168,8
5'-TGTXAC-3'	1879,6	1880,3
5'-GTXAC-3'	1575,5	1575,9
5'-TXAC-3'	1246,5	1245,4

B

5'-CTGTXAC-3'

3'-exonuclease	MALDI-TOF/MS Cald. mass	Exp. mass
5'-CTGTXAC-3'	2169,6	2168,1
5'-CTGTXA-3'	1879,6	1879,5
5'-CTGTX-3'	1566,5	1566,6

C

5'-CT
5'-GTXAC-3'

5'-exonuclease	MALDI-TOF/MS Cald. mass	Exp. mass
5'-CT / 5'-GTXAC-3'	2169,6	2168,8
5'-CT / 5'-TXAC-3'	1839,6	1839,8
5'-T / 5'-GTXAC-3'	1879,6	1879,6
5'-T / 5'-TXAC-3'	1550,5	1550,0

D

5'-CT
5'-GTXAC-3'

3'-exonuclease	MALDI-TOF/MS Cald. mass	Exp. mass
5'-CT / 5'-GTXAC-3'	2169,6	2168,8
5'-CT / 5'-GTXA-3'	1879,6	1879,5
5'-CT / 5'-GTX-3'	1566,5	1566,0

Conclusions

We have developed a single automated chemical synthesis of a branched oligonucleotide bearing a biologically relevant lesion. For this, we have successfully prepared a phosphoramidite derivative bearing an orthogonally protected group at the branching point. We have shown that this phosphoramidite building block could be incorporated into oligonucleotides of different lengths and could be used, after deprotection, to produce branched oligonucleotides. Preparation of longer oligonucleotides will be performed in order to use them for biological studies including the study of repair and mutagenicity of the dCyd341 interstrand crosslink.

Experimental

View Article Online
DOI: 10.1039/C9OB01021A

Instrumentation. Oligonucleotides were prepared on a 392 DNA synthesizer (Applied Biosystems Inc, Palo Alto, CA). The HPLC-MS/MS analyses were performed with a triple quadrupole mass spectrometer TSQ Quantum Ultra from Thermo Scientific equipped with a Thermo Accela Pump, a Accela autosampler and a Accela photodiode array detector. The data were processed using Xcalibur software. HPLC purification was performed using a Agilent 1200 Series LC system. 1H NMR spectra were recorded on a Brucker Avance DMX 200 at 295 K in 400 µL of either CCl_3D. The MALDI-TOF/MS mass spectra were obtained in the negative mode on a time-of-flight Microflex mass spectrometer (Bruker), equipped with a 337-nm nitrogen laser and pulsed delay source extraction. The matrix was prepared by dissolving 3-hydroxypicolinic acid in 10 mM ammonium citrate buffer. Prior to MALDI-TOF mass analysis, oligonucleotide solutions were purified and concentrated on ZipTip pipette tips (Millipore). A mixture of purified DNA sample (10 pmol; 1 µl) was added to matrix (1 µl) and spotted on a polished stainless target plate using the dried droplet method. Spectra were calibrated using reference oligonucleotides of known masses.

Synthesis of dCyd341 phosphoramidite

Furfuryllevulinate (4). To a solution of furfuryl acohol (0.5 ml, 5.8 mmol), DCC (3.0 g, 14.6 mmol) and DMAP (70 mg, 0.6 mmol) in THF (100 ml), 20 ml of a THF solution of levulinic acid (1.2 ml, 11.6 mmol) was added. After 3 h with stirring at room temperature the mixture was carefully filtered and the supernatant was evaporated. The resulting residue was dissolved in DCM, and washed successively with saturated sodium bicarbonate, water and brine. Finally it was dried over Na_2SO_4 and concentrated at reduced pressure to give compound **4** in quantitative yields. ^{1}H NMR (Fig S15) (200 MHz, Cl_3CD)δ (ppm) 7.39 (dd, J = 1.8, 0.9 Hz, 1H), 6.38 (d, J = 3.2Hz, 1H), 6.33 (dd, J = 3.2, 1.8 Hz, 1H), 5.03 (s, 2H), 2.73-2.69 (m, 2H),2.60-2.57 (m, 2H) 2.15 (s, 3H). MS m/z 235 ($M + K^+$), 219 ($M + Na^+$), 197 ($M+H^+$).

Trans 2,5-dioxo-pent-3-enyl levulinate (5). Compound **4** (1.13 g, 5.8 mmol) was dissolved in a mixture of acetone:water 10:1 (25 ml) and placed in a cooling bath at – 20 °C. $NaHCO_3$ (1.00 g, 11.9 mmol) was added to the cooled solution. A solution of N-bromosuccinimide (NBS) (1.22 g, 6.85 mmol) in the same solvent (16 ml) was then added dropwise with stirring. After the addition was complete the reaction was stirred for 1 h at -20 °C. The reaction was then quenched by addition of pyridine (2 ml, 20 mmol) and removed from the cooling bath. It was then allowed to warm at room temperature and stirred for 2 additional hours. The mixture was poured onto EtAcO and extracted several times with HCl 5%, H_2O, saturated sodium bicarbonate and brine. Finally, it was dried over Na_2SO_4 and concentrated at reduced pressure to give a crude mixture with 75 % yield of compound **5** (measured by ^{1}H-NMR). The crude mixture was subsequently used without further purification. ^{1}H NMR (Fig. S16)(200 MHz, Cl_3CD) δ (ppm) 9.77 (d, J = 7.0 Hz, 1H), 7.40 (m, 0.4H), 7.04 (d, J = 16.2 Hz, 1H), 6.83 (dd, J = 16.2, 7.0 Hz, 1H), 6.38-6.34 (m,

0.8H), 5.03 (s, 0.8H), 4.90 (s, 2H), 2.87 – 2.48 (m, 5.6 H), 2.17 (s, 3H), 2.15 (s, 1.2 H).

6-(5´-O-(DMTr)-3´-O-TBDMS-2´-D-deoxyribofuranosyl)-2-hydroxy-3(3-hydroxy-2-oxopropyl)-2,6-dihydro-imidazo[1,2-c]-pyrimidin-5(3H)-one (5´-O-DMT-3´-O-TBDMS-dCyd341) (6). To a stirring solution of the protected deoxycytidine prepared as previously described[25] (0.6 g, 0.9 mmol) in DCM (15 ml) a solution of **5** (1.20 g .4.5 mmol) in the same solvent (10 ml) was added dropwise under Ar atmosphere and the reaction mixture was then stirred overnight at room temperature. Afterwards, the reaction mixture was washed successively with saturated sodium bicarbonate, water and brine. Finally it was dried over Na_2SO_4 and concentrated at reduced pressure. The resulting oil was purified by flash column chromatography (DCM:MeOH 100:3) to give compound **6** as brown oil (93 % yield).^{1}H NMR (Fig S17) (200 MHz, Cl_3CD) δ (ppm) 7.32 – 7.29 (m, 10H), 6.84 – 6.79 (m, 4H), 6.35 – 6.16 (m, 2H), 5.53 – 5.50 (m, 1H), 4.77 – 4.68 (m, 1H), 4.57 – 4.48 (m, 1H), 4.27-4.18 (m, 2H), 3.97 – 3.92 (m, 1H), 3.82 (s, 6H), 3.43 – 3.31 (m, 2H), 2.90 – 2.78 (m, 5H), 2.22 (s, 3H), 0.87 (s, 9H), 0.07 – 0.01 (m, 6H).ESI MS m/z 856 ($M+H^+$), 324 (BH^{2+}), 303 ($DMTO^+$).

6-(5´-O-(DMTr)-3´-O-TBDMS-2´-D-deoxyribofuranosyl)-2-O-Acetyl-3(3-hydroxy-2-oxopropyl)-2,6-dihydro-imidazo[1,2-c]-pyrimidin-5(3H)-one (5´-O-DMT-3´-O-TBDMS-2-O-Ac-dCyd341) (7). To a solution of compound **6** (0.74 g, 0.9 mmol) in pyridine (6 ml) acetic anhydride (0.16 ml, 1.8 mmol) and catalytic amounts of DMAP were added under Ar atmosphere. The mixture was stirred overnight at room temperature and then concentrated under reduced pressure to dryness. The residue was dissolved in DCM and then was washed successively with saturated sodium bicarbonate, water and brine. Finally, it was dried over Na_2SO_4 and concentrated at reduced pressure. The resulting oil was purified by flash column chromatography (DCM:MeOH 100:1) to give compound **12** (30 % yield) and the mixture of **7** : **8** (63% yield, 3:1 by ^{1}H-NMR). Mixture of compounds **7** and **8**. ^{1}H NMR (Fig S18) (200 MHz, Cl_3CD) δ (ppm) 7.72-7.59 (m, 1H (7 and 8)), 7.42– 7.29 (m, 10H), 7.19 (s, 0.2H (8)), 6.89 – 6.85 (m, 4H), 6.52 – 6.45 (m, 0.8H(7)), 6.45 – 6.33 (m, 1H (7 and 8)), 6.26 (m, 0.2H (8)), 5.58 (m, 0.8H (7)), 4.90 (s, 0.4H (8)), 4.70 – 3.91(several peaks, m, 5.4H(7 and 8)), 3.83 (s, 6H), 3.56 – 3.36 (m, 2H(7 and 8)), 2.84 – 2.58 (m, 6H(7 and 8)), 2.21 – 2.07 (several singlets, 5.4H (7 and 8)), 0.85 (s, 9H (7 and 8)), 0.12 – 0.01 (m, 6H (7 and 8)). ESI MS m/z 898 ($M+H^+$), 838 ($M+H^+$), 366 (BH^{2+}), 306 (BH^{2+}), 303 ($DMTO^+$). Compound **8** (enriched for characterization, Fig. S19) ^{1}H NMR (200 MHz, Cl_3CD) δ (ppm) 7.70 (d, J = 7.9 Hz, 1H), 7.40 – 7.30 (m, 10H) 7.21 (s, 1H), 6.92 – 6.86 (m, 4H), 6.45 (t, J = 5.8 Hz, 1H), 6.25 (d, J = 7.9 Hz, 1H), 4.91 (s, 2H), 4.60 (q, J = 5.1 Hz, 1H), 4.19 (s, 2H), 4.00 – 3.96 (m, 1H), 3.83 (s, 6H), 3.56 (dd, J = 10.8, 2.8 Hz, 1H), 3.37 (dd, J = 10.8, 2.9 Hz, 1H), 2.95-2.71 (m, 5H), 2.48-2.43 (m, 1H), 2.24 (s, 3H), 0.86 (s, 9H), 0.06 (s, 3H), 0.00 (s, 3H). ESI MS m/z 838 ($M+H^+$), 306 (BH^{2+}), 303 ($DMTO^+$). Compound **12**. ^{1}H NMR (Fig S20) (200 MHz, Cl_3CD) δ (ppm) 7.81 (d, J = 7.9 Hz, 1H), 7.64 (s, 1H), 7.62 (s, 1H), 7.46 -7.31 (m, 10H), 6.89 – 6.85 (m, 4H), 6.49 (t, J = 5.7 Hz, 1H), 6.31 (d, J = 7.9 Hz, 1H), 4.89 (s, 2H), 4.59 (q, J = 5.2 Hz, 1H), 4.12 – 4.02 (m, 1H), 3.82 (s, 6H), 3.59 (dd, J = 10.8, 2.6 Hz, 1H), 3.55 (dd, J = 10.8, 2.7 Hz, 1H), 2.94-2.77 (m, 3H), 2.69-2.63 (m, 2H), 2.54-2.41 (m, 1H), 2.32 (s, 3H), 2.22 (s, 3H), 0.86 (s, 9H), 0.05 (s, 3H), 0.00 (s, 3H). ESI MS m/z 880 ($M+H^+$), 348 (BH^{2+}), 303 ($DMTO^+$).

6-(5´-O-(DMTr)-2´-D-deoxyribofuranosyl)-2-O-Acetyl-3(3-hydroxy-2-oxopropyl)-2,6-dihydro-imidazo[1,2-c]-pyrimidin-5(3H)-one (5´-O-DMT-2-O-Ac-dCyd341) (9). A mixture of **7** and **8** (0.5 g, 0.6 mmol) in THF (30 mL) and Bu_4NF (1.2 mmol, 1.2 mL in 1 M THF solution) was stirred under Ar atmosphere for 1 h at 0°C. The reaction mixture was then diluted with ether (60 mL) and washed with water, brine, and dried over Na_2SO_4. After solvent was evaporated, the residue was purified by column chromatography (DCM:MeOH:Et_3N 100:1:0.5) to give the mixture of compound **9** and **10** in 69 % yield.^{1}H NMR (Fig S21) (200 MHz, Cl_3CD) δ (ppm) 7.77-7.43 (m, 1H), 7.36– 7.29 (m, 10H), 7.16 (s, 0.5H), 6.98 – 6.85 (m, 4H), 6.50 – 6.26 (m, 2H), 5.34 (s, 1H),4.90 (s, 1H), 4.70 – 4.58(m, 1.5H), 4.43-4.04 (m, 3.5H), 3.82 (s, 6H), 3.58 – 3.44 (m, 2H), 2.88 – 2.81 (m, 3.5H), 2.77 – 2.33 (m, 4H), 2.23-1.93 (several singlets, 4H). ESI MS m/z 822 ($M + K^+$), 806 ($M + Na^+$), 784 ($M+H^+$), 762 ($M' + K^+$), 746 ($M' + Na^+$), 724 ($M'+H^+$), 366 (BH^{2+}), 306 ($BH^{2+'}$), 303 ($DMTO^+$).

6-(3'-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5´-O-(DMTr)-2´-D-deoxyribo-furanosyl)-2-O-Acetyl-3(3-hydroxy-2-oxopropyl)-2,6-dihydro-imidazo[1,2-c]-pyrimidin-5(3H)-one(5´-O-DMT-2-O-Ac-dCyd341phosphoramidite) (11). 2-Cyanoethyl N,N-diisopropylchlorophosphoramidite (0.23 mL, 1.05 mmol) was added to a solution of **9** and **10** (0.32g, 0.42 mmol) and DIPEA (0.33 mL, 1.8 mmol) in DCM (15 mL). The reaction mixture was stirred at room temperature for 80 min and then 5% $NaHCO_3$ solution and EtOAc were added. The organic layer was separated, dried over Na_2SO_4, and concentrated under reduced pressure. Purification by flash chromatography (EtOAc : Et_3N 100:0.5) provided **11** (160 mg, 40 %) and **13** (145 mg, 31%). Compound **11**.^{1}H NMR (Fig S22) (200 MHz, Cl_3CD) δ (ppm) 7.62 (dd, J = 8.0, 10.0 Hz, 1H), 7.43 – 7.30 (m, 10H), 6.91 – 6.85 (m, 4H), 6.53 – 6.22 (m, 2H), 5.57 – 5.34 (m, 1H), 4.91 – 4.67 (m, 2H), 4.47 – 4.15 (m, 3H), 3.84 (s, 6H), 3.70 – 3.45 (m, 5H), 2.88 – 2.32 (m, 10H), 2.21 – 2.08 (several singlets, 3H), 1.63 (s, 3H) 1.33 – 1.20 (several singlets, 12H). ESI MS m/z 984 ($M+H^+$), 619 ($M-BH^+$), 366 (BH^{2+}), 303 ($DMTO^+$). Compound **13**.^{1}H NMR (Fig S23) (200 MHz, Cl_3CD) δ (ppm) 7.98 (s, 1H), 7.64 (dd, J = 14.0, 8.0 Hz, 1H), 7.48-7.30 (m, 10H), 7.13 (s, 1H), 6.91-6.85 (m, 4H), 6.57 (m, 1H), 6.27 (dd, J = 7.8, 1.8 Hz, 1H), 5.00 (d, J = 13.1 Hz, 1H), 4.84 (d, J = 13.1 Hz, 1H), 4.75-4.67 (m, 1H), 4.23-4.18 (m, 1H), 3.99-3.88 (m, 3H), 3.83 (m, 8H). 3.76-3.47 (m, 7H), 2.88-2.82 (m, 2H), 2.73-2.64 (m, 6H), 2.25 (s, 3H), 1.26-1.2 (m(several singlets), 24H). ESI MS m/z 1162 ($M + K^+$), 1146 ($M + Na^+$), 1124 ($M+H^+$), 506 (BH2+), 303 (DMTO+).

Branched oligonucleotide synthesis

Oligonucleotides bearing the modified phosphoramidite were prepared using fast deprotection phosphoramidite chemistry. The standard 1 µmole DNA cycle with coupling times extended to 15 min for the condensation of the modified phosphoramidite. Then, a separate end-capping cycle (consisting of detritylation and acetylation) was used to block further elongation at the 5'-end. The

support was finally manually treated with a 0.5 M solution of hydrazine hydrate in pyridine : acetic acid (1:1) for 15 min to remove the levulinyl protecting group. Synthesis was then continued in the normal fashion from the free base hydroxyl group at the branching point, using again traditional phosphoramidite chemistry with retention of the 5' terminal DMT group (DMT-ON). The oligonucleotides were removed from the solid support by treatment with concentrated ammonia at room temperature for 4 h. Solvents were then removed by evaporation under reduced pressure.

HPLC-UV purification and analysis

The crude oligonucleotide mixtures were purified and deprotected on-line by reverse phase HPLC[25]. Purification was achieved using a Hamilton PRP3 polymeric phase column, 10 µm, 305 × 7.0 mm and a mixture of acetonitrile and 10 mM TEAA buffer as the eluents [100% TEAA (5 min), linear gradient from 0 to 8% of acetonitrile (10 min), then isocratic TEAA-acetonitrile (92/8) v/v (10 min). Then, in order to remove the DMTr group, column was eluted under isocratic condition using 100% TFA (1%) for 10 min and finally a gradient from 0 to 10% acetonitrile (40 min) with a flow rate of 2.5 mL min^{-1} was used to elute the fully unprotected oligonucleotide. UV detection was set at 260 nm. Fractions were collected and the solvent was evaporated under reduced pressure. Oligonucleotides were characterized by MALDI-TOF/MS and by HPLC-UV. For oligonucleotides analyses, chromatographic separations were achieved using a mixture of acetonitrile and 10 mM TEAA as the eluents [100% 10 mM TEAA (4 min), linear gradient from 0 to 100% of acetonitrile (26 min)].

Oligonucleotide digestion and HPLC-MS/MS analysis.

Incorporation of dCyd341 in oligonucleotides was confirmed by HPLC-MS/MS following enzymatic digestion as described in detail elsewhere.[26] Chromatographic separations were achieved using a ODB uptisphere 3 µm 150*2.1 mm column from Interchim under linear gradient conditions as previously reported for measuring other DNA lesions.[27,28] All nucleosides were detected using a UV detector set at 260 nm. Multiple reaction monitoring mode was applied for the detection of dCyd341 lesion by using transitions 342→226, 342→168, and 342→148.

Exonuclease digestion

DOI: 10.1039/C9OB01031A

Branched and non-branched oligonucleotides (200pmol) were dissolved in 15.5 µL of water and 4 µL of 500 mM ammonium citrate was added to the solution. For the 3'->5' digestion, 0.5 µL phosphodiesterase I (crotalus venom) (0.03 units/µL) was added to the sample solution, whereas for the 5'->3' digestion, 1 µL of phosphodiesterase II (bovine spleen) (0.01 units µ/L) was added. The solutions were incubated at 37 °C. Aliquots of the digested solutions (1 µL) were removed at fixed periods of time and dissolved in 9 µL of water and placed on ice to interrupt the digestion reaction. Aliquots (1 – 2 µL) of the samples and the trihydroxyacetophenone matrix solution were mixed and immediately spotted on a MALDI-TOF/MS target and air-dried before analysis. The mass spectrometer was operated in the negative linear mode.

Conflicts of interest

There are no conflicts to declare

Acknowledgements

This work was performed within the framework of the LABEX PRIMES (ANR-11-LABX-0063) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR). The authors are grateful to the Agence Nationale de la Recherche grant ANR-09-PIRI-0022, Radiobiology program of CEA, CEA Eurotalents program (to MTP), Labex ARCANE and CBH-EUR-GS (ANR-17-EURE-0003) for financial support.

Notes and references

(1) Takayanagi, R.; Inoguchi, T.; Ohnaka, K. *J Clin Biochem Nutr* **2011**, *48*, 72.

(2) Elahi, M. M.; Kong, Y. X.; Matata, B. M. *Oxid Med Cell Longev* **2009**, *2*, 259.

(3) Devine, M. J.; Plun-Favreau, H.; Wood, N. W. *Nat Rev Cancer* **2011**, *11*, 812.

(4) Kryston, T. B.; Georgiev, A. B.; Pissis, P.; Georgakilas, A. G. *Mutat Res* **2011**, *711*, 193.

(5) Cadet, J.; Ravanat, J.-L.; Taverna-Porro, M.; Menoni, H.; Angelov, D. *Cancer letters* **2012**, *327*, 5.

(6) Eccles, L. J.; O'Neill, P.; Lomax, M. E. *Mutat. Res.* **2011**, *711*, 134.

(7) Deans, A. J.; West, S. C. *Nature Reviews Cancer* **2011**, *11*, 467.

(8) Mitchell, J. R.; Hoeijmakers, J. H. J.; Niedernhofer, L. J. *Current Opinion in Cell Biology* **2003**, *15*, 232.

(9) Lawley, P. D.; Phillips, D. H. *Mutat Res* **1996**, *355*, 13.

(10) Noll, D. M.; Mason, T. M.; Miller, P. S. *Chem Rev* **2006**, *106*, 277.

(11) Niedernhofer, L. J.; Daniels, J. S.; Rouzer, C. A.; Greene, R. E.; Marnett, L. J. *J Biol Chem* **2003**, *278*, 31426.

(12) Ravanat, J. L.; Breton, J.; Douki, T.; Gasparutto, D.; Grand, A.; Rachidi, W.; Sauvaigo, S. *Br. J. Radiol.* **2014**, *87*, 20130715.

(13) Regulus, P.; Duroux, B.; Bayle, P.-A.; Favier, A.; Cadet, J.; Ravanat, J.-L. *Proc. Natl. Acad. Sci., U.S.A.* **2007**, *104*, 14032.

(14) Sczepanski, J. T.; Jacobs, A. C.; Greenberg, M. M. *J. Am. Chem. Soc.* **2008**, *130*, 9646.

(15) Warren, A. J.; Hamilton, J. W. *Chem Res Toxicol* **1996**, *9*, 1063.

View Article Online
DOI: 10.1039/C9OB01021A

(16) Fischhaber, P. L.; Gall, A. S.; Duncan, J. A.; Hopkins, P. B. *Cancer Res* **1999**, *59*, 4363.
(17) Alzeer, J.; Scharer, O. D. *Nucleic Acids Res* **2006**, *34*, 4458.
(18) Wilds, C. J.; Xu, F.; Noronha, A. M. *Chem Res Toxicol* **2008**, *21*, 686.
(19) Ye, M.; Guillaume, J.; Liu, Y.; Sha, R.; Wang, R.; Seeman, N. C.; Canary, J. W. *Chem Sci* **2013**, *4*, 1319.
(20) Gamboa Varela, J.; Gates, K. S. *Angew Chem Int Ed Engl* **2015**, *54*, 7666.
(21) Stevens, K.; Madder, A. *Nucleic Acids Res* **2009**, *37*, 1555.
(22) Katolik, A.; Johnsson, R.; Montemayor, E.; Lackey, J. G.; Hart, P. J.; Damha, M. J. *J Org Chem* **2014**, *79*, 963.
(23) Ghosh, S.; Greenberg, M. M. *J Org Chem* **2014**, *79*, 5948.
(24) Chandra, M.; Keller, S.; Gloeckner, C.; Bornemann, B.; Marx, A. *Chemistry* **2007**, *13*, 3558.
(25) Romieu, A.; Gasparutto, D.; Molko, D.; Cadet, J. *Tet. Lett.* **1997**, *38*, 1531.
(26) Ravanat, J.-L.; Douki, T.; Duez, P.; Gremaud, E.; Herbert, K.; Hofer, T.; Lasserre, L.; Saint-Pierre, C.; Favier, A.; Cadet, J. *Carcinogenesis* **2002**, *23*, 1911.
(27) Frelon, S.; Douki, T.; Ravanat, J.-L.; Pouget, J.-P.; Tornabene, C.; Cadet, J. *Chem. Res. Toxicol.* **2000**, *13*, 1002.
(28) Ravanat, J.-L. *Free Radic. Res.* **2012**, *46*, 479.