Characterization of Byproducts from Chemical Syntheses of Oligonucleotides Containing 1-Methyladenine and 3-Methylcytosine

Qi Tang, ‡✉ Ang Cai, ‡✉ Ke Bian, ‡ Fangyi Chen, ‡✉ James C. Delaney, ‡ Sravani Adusumalli, ‡ Alvin C. Bach, II, ‡ Fatemeh Akhlaghi, ‡ Bongsup P. Cho, ‡ and Deyu Li* ‡

†Department of Biomedical and Pharmaceutical Sciences, College of Pharmacy, University of Rhode Island, 7 Greenhouse Road, Kingston, Rhode Island 02881, United States
‡Visterra Inc., One Kendall Square, Cambridge, Massachusetts 02139, United States

Supporting Information

ABSTRACT: Oligonucleotides serve as important tools for biological, chemical, and medical research. The preparation of oligonucleotides through automated solid-phase synthesis is well-established. However, identification of byproducts generated from DNA synthesis, especially from oligonucleotides containing site-specific modifications, is sometimes challenging. Typical high-performance liquid chromatography (HPLC), mass spectrometry (MS), and gel electrophoresis methods alone are not sufficient for characterizing unexpected byproducts, especially for those having identical or very similar molecular weight (MW) to the products. We used a rigorous quality control procedure to characterize byproducts generated during oligonucleotide syntheses: (1) purify oligonucleotides by different HPLC systems; (2) determine exact MW by high-resolution MS; (3) locate modification position by MS/MS or exonuclease digestion with matrix-assisted laser desorption ionization-time of flight analysis; and (4) conduct, where applicable, enzymatic assays. We applied these steps to characterize byproducts in the syntheses of oligonucleotides containing biologically important methyl DNA adducts 1-methyladenine (m1A) and 3-methylcytosine (m3C). In m1A synthesis, we differentiated a regiosomeric byproduct 6-methyladenine, which possesses a MW identical to uncharged m1A. As for m3C, we identified a deamination byproduct 3-methyluracil, which is only 1 Da greater than uncharged m3C in the ~4900 Da context. The detection of these byproducts would be very challenging if the abovementioned procedure was not adopted.

INTRODUCTION

Oligonucleotides synthesized chemically are widely used as drugs and research tools in biology, chemistry, and medicine. Solid-phase synthesis of DNA and RNA oligonucleotides has been well-developed, and the automated phosphoramidite-based chemical process has become highly efficient.1−3 Phosphoramidites of regular and some modified nucleotides are commercially available, and certain oligonucleotide strands could be readily obtained from commercial sources.4−6 Besides the great development in the synthesis of oligonucleotides, the differentiation of byproducts from the product oligonucleotides, especially on those containing site-specifically modified structures, is sometimes ignored by the end users. If those byproducts or small impurities were not identified and removed, it could have devastating consequences for the corresponding biological assays and medical treatments.6−9 For certain instances, it is challenging to identify some byproducts generated in the synthesis and deprotection steps, especially the byproduct that has a molecular weight (MW) identical or very similar to the desired product oligonucleotide.

We have synthesized various oligonucleotides containing modified structures in the past, focusing on alkyl or aryl DNA adducts, by using solid- and solution-phase DNA synthesis.10−14 In the syntheses of oligonucleotides containing 1-methyladenine (m1A) and 3-methylcytosine (m3C), we observed byproducts in a neutral context that are either regiosomerical (identical MW) to the product or have 1 Da in MW greater than the product in a ~4900 Da context (Figure 1a). m1A and m3C are formed by exogenous and endogenous alkylating agents mainly in single-stranded DNA and have been detected both in vitro15−21 and in vivo.20,22−29 Both adducts are cytotoxic and block DNA replication and are the best substrates for the AlkB family DNA repair enzymes (Figure 1b).30−34 m3C has also been proposed as an epigenetic biomarker for cancer.35 Most of our DNA syntheses successfully provided target oligonucleotides. However, some oligonucleotides contained side reaction contaminants generated during the synthesis or deprotection steps, requiring further purification and identification. To that end, we applied a rigorous quality control process which entails the following steps: (i) synthesize an adduct-containing oligonucleotide from the corresponding phosphoramidite; (ii) purify the product oligonucleotide by both reverse-phase and anion-exchange high-performance liquid chromatography (HPLC); (iii) measure the exact MW of the oligonucleotide by high-resolution mass spectrometry.
(HR-MS), certain impurities can also be detected by HR-MS; (iv) determine the modification position by either MS/MS or exonuclease digestion with matrix assisted laser desorption ionization-time of flight (MALDI-TOF) analysis; and (v) test the biological activity of the adduct-containing oligonucleotide by appropriate enzymatic assays. We are assuming that the modified nucleotide has been extensively investigated for its stability under solid-phase chemical synthesis and deprotection beforehand; however, the strategy outlined above should confirm this. Nuclear magnetic resonance (NMR; $^1$H, $^{13}$C, and $^{31}$P) is a powerful tool for structural integrity studies of the phosphoramidite on the milligram scale. However, NMR becomes impractical for characterizing minor impurities within the modified structures in synthetic oligonucleotides because of increasing spectral complexity, as well as the final amount of material available (e.g., on a micro-mole or less scale). Below, we report the detailed characterization of byproducts in the syntheses of oligonucleotides containing m1A and m3C. The characterization and separation of those byproducts provide confidence in the quality of oligonucleotides used in further biological experiments.

**RESULTS**

**Identification of the 6-Methyladenine Byproduct from m1A Synthesis.** We were trying to incorporate m1A into an oligonucleotide as a substrate for the AlkB repair reaction. A 16mer oligonucleotide containing m1A was prepared by using the commercially available phosphoramidite of m1A (Figure 1a). The final product was deprotected under standard conditions by treating the crude oligonucleotide with ammonium hydroxide at 80 °C for 3 h. The oligonucleotide was then tested by both reverse-phase and anion-exchange HPLC (see the Experimental Section for detailed conditions). The resulting chromatograms showed a single peak under both conditions (retention time 10.2 min in Figure 2a and 4.0 min in Figure S22). HR electrospray ionization-time of flight (ESI-TOF) MS analysis of the sample exhibited $m/z$ at 1224.715 at its −4 charge state, which is in good agreement with the theoretical $m/z$ 1224.711 expected of the product oligonucleotide (Figures 3a and S1 and Table 1).

Collision-induced dissociation (CID) MS/MS analysis was used to determine the location of the alkyl adduct in the synthesis of m1A-containing 16mer oligonucleotide (Figure 4). The MS/MS results, presented in detail in the Supporting Information (Figure S5 and Table S1), confirmed that modification occurred at the eighth position from the 3′ end of the oligonucleotide, and the modified base had the same MW as m1A. Analysis of 3′ end fragmentations indicated that $w_7$ to $w_8$ ions showed only standard DNA sequences without methyl modification (Table S1). These results suggested that the modification is located between $w_8$ and $w_7$, which is the eighth position from the 3′ end of the 16mer oligonucleotide. From the S′ fragmentations, we observed the methyl modification in ions from (a15-C) to (a10-G), but no such modification in ions from (a9-X) to (a2-A) (Figures 4 and S9 and Table S1). These results suggested that the modification is located between (a10-G) and (a9-X), which is the ninth position from the S′ end (Figure 4). The fragmentation patterns from both the S′ and 3′ ends are consistent with the proposed m1A-containing DNA sequence (Figure 4).

To further confirm the location of the modified base, we performed enzyme digestion coupled with MALDI-TOF MS analysis. The procedure has been widely used to confirm the location of lesion positions by determining the mass changes after partial digestion from the 3′ end by the exonuclease snake venom phosphodiesterase (SVP).

Figure 5 shows the positive ion MALDI-TOF MS spectra of the 3′ to S′ SVP exonuclease digestion of the product oligonucleotide at different time intervals. The $m/z$ of 4904.2 at 0 min (before digestion, as a control) indicated that the 16mer DNA sequence contained a methylated DNA base ([M +H]$^+$ theoretical $m/z$ = 4903.9 Da; herein theoretical $m/z$ = 8212).
values are shown in parentheses after the observed m/z values). After 1 min SVP digestion, while the product ion disappeared, three new lower masses appeared at m/z = 4615.7 (4614.8), 4326.7 (4325.8), and 4022.7 (4021.7), which correspond to the 15, 14, and 13mer fragments generated from 3′ cleavage, respectively (Figure 5). The signal intensity of three new peaks was significantly increased with the disappearance of the original product oligonucleotide ion at m/z = 4904.2.

Further digestion resulted in a number of smaller fragments. In the 3 min digestion (Figures 5 and S17), the signal of m/z = 3075.1 (3074.6) matches the sequence for the fragment that was cleaved one base before the lesion. The m/z = 2746.4 (2745.5, containing a methyl modification) represents the digestion occurring on the 3′ side of the modified adenine position (Figure 5), which also persisted in the 5 min digestion spectrum. The signal at 2419.1 (2418.5, no methyl modification) in the 3 and 5 min digestions represents the fragment generated after liberation of deoxymethyladenosine 5′ monophosphate. Digestions practically stopped at m/z = 1223.6/8 (1222.3), which represents the m/z of a 4mer nucleotide of the 5′ end. Taken together, these results confirm that the modified structure is a methylated adenine and it is at the ninth position of the proposed oligonucleotide from the 5′ to 3′ direction.

The chromatographic and mass spectral evidence presented above indicates that the product oligonucleotide could have the MW expected for a 16mer containing the m1A adduct in the correct position. We then tested the repair of this product by the AlkB protein, which is an enzyme that repairs various alkylated DNA adducts (Figure 1b).32,34 The result showed the repair efficiency of this oligonucleotide by AlkB was very low, even when 2.5 μM of AlkB was incubated with the 5.0 μM product oligonucleotide (Figures 6b and S24). This observation was in contrast to the report that m1A is a good substrate of AlkB (e.g. 0.2 μM of AlkB is able to repair 5.0 μM m1A completely in 1 h).10 We suspected that the structure in the product oligonucleotide may not be m1A, even though it had a MW identical to m1A. The reverse-phase HPLC results also showed an anomalous phenomenon. For example, the positively charged m1A nucleobase35 should be more polar than the natural base adenine and thus have a shorter retention time on a C18 column. However, the opposite was observed: adenine has a shorter retention time than the product (Figure 2a-c).

It has been reported that 6-methyladenine (m6A) could be generated from a Dimroth rearrangement of m1A under basic conditions (Figure 1a and Table 1).34,45 We therefore conducted a new batch synthesis of m1A under a milder deprotection condition (25 °C for 16 h). The product oligonucleotide was eluted as a single peak in both HPLC columns (Figure 2b for C18 and Figure 6d for anion-exchange). The HPLC retention time in reverse-phase HPLC exhibited a shorter retention time (8.5 min, Figure 2b) than the unmodified control adenine (9.0 min, Figure 2c) or the previous product (10.2 min, m6A, Figure 2a), which is in good agreement with the theory that m1A should be more polar than both A and m6A.

The MW of the new product was characterized by HR-MS and had the correct m/z of m1A at 1224.715 (1224.711) (Figures 3b and S2 and Table 1). The MS/MS results (e.g., the w2, w3, a2X, and a3G ions, Figures 4, S6, S11, and S12, and Table S2) demonstrated that the modification is at the ninth position from the 5′ end and it is a methylated adenine. The exonuclease digestion with MALDI-TOF analyses (Figures S18 and S21a) also confirmed the location and identity of the adduct. For example, the 2418.1 (before the adduct) and 2418.8 (after the adduct) peaks at 3 min digestion in Figure S18 were present. Finally, incubation with AlkB showed an

| lesion or base | MW (calculated) of neutral species | m/z (calculated) −4 charge peak | m/z (observed) −4 charge peak |
|---------------|-----------------------------------|-------------------------------|-----------------------------|
| 16mer m6A     | 4902.877                          | 1224.711                      | 1224.715                    |
| 16mer m1A     | 4902.877                          | 1224.711                      | 1224.715                    |
| 16mer m3C     | 4878.866                          | 1218.709                      | 1218.703                    |
| 16mer m3U     | 4879.850                          | 1218.955                      | 1218.957                    |

For m1A and m3C syntheses, the sequence of the 16mer was 5′-GAAGACCTXGGCCTCC-3′, where X indicates the position of the modified bases.
excellent repair efficiency against the new product oligonucleotide:
5.0 μM adduct was mostly repaired in 1 h with the 0.2 μM
enzyme (Figure 6e).10 Taken together, these results confirmed
the newly synthesized product as a 16mer oligonucleotide
containing m1A at the ninth position from the 5′ end (Figure
1a).

Identification of the m3U Byproduct from m3C
Synthesis. We also wanted to synthesize an oligonucleotide
containing m3C as the substrate for the AlkB family DNA
repair reactions within the same sequence context as for the
m1A oligonucleotide (Figure 1a).34,39 After automated syn-
thesis and standard deprotection (80 °C for 3 h), the HPLC
chromatogram from a C18 column showed two peaks with similar intensities (9.5 and 11.6 min, Figure 2d), which were readily separated (Figure 2e,f). HR-MS analysis showed that the early eluting peak (9.5 min, Figure 2e) had an m/z value of 1218.703 at its −4 charge state (Figures 3c and S3), which is in agreement with that of the m/z of the target m3C-containing oligonucleotide (theoretical m/z 1218.709, Table 1). The second species at 11.6 min (Figure 2f) showed an m/z of 1218.957 at its −4 charge state (Figures 3d and S4). Compared to the 9.5 min peak, the m/z value of the second species was 0.254 unit higher at the −4 charge state, indicating that the MW (4879.850 Da) of the second species is 1.016 Da higher than the first one. This is consistent with the m/z value of an oligonucleotide containing 3-methyluracil (m3U) at its −4 charge state (1218.955, Table 1). m3U could potentially be generated from deamination of m3C under basic conditions (Figure 1a). The MS/MS results of the two oligonucleotides (e.g., the w7,w8,a9-X, and a10-G ions, Figures 4, S7, S8, and S13 to S16, Tables S3, and S4) demonstrated that the modifications are at the ninth position from the 5′ end. The exonuclease digestion with MALDI-TOF analyses (Figures S19, S20, and S21b) also confirmed the location and identity of the adducts. At 7 min digestion, signature peaks were 2722.2 (before the adduct) and 2723.3 (before the adduct) and 2418.7 (after the adduct) (Figure 6h). This was not the case for m3U-containing digestion with MALDI-TOF analysis. For HPLC analysis of modified oligonucleotides, we used a combination of two different separation systems, such as reverse-phase (e.g. C18) and anion-exchange chromatography. For example, it is not easy to find a suitable condition for fully separating 16mer-containing m1A and A with C18 columns (Figure 2b,c). However, it is relatively easy to separate them with an anion-exchange column because m1A is positively charged at neutral pH but A is not (Figures 1a, 6d and 6f). On the other hand, it is not easy to distinguish m6A and A under anion-exchange conditions (Figures S22 and S23) because of their neutrality at pH 7.0, but it is possible to separate them on the C18 column (Figure 2a,c).

The MW of the oligonucleotide can be determined by HR-MS, and the location of the modification can be identified by MS/MS or exonuclease digestion with MALDI-TOF analysis. To further confirm the identity of the oligonucleotide from a biological activity perspective, we recommend running enzymatic reaction on the product. In this work, we used SVP exonuclease to digest all oligonucleotides formed in the synthesis. A specific enzyme for certain types of modification could greatly help elucidate the structure of modification. Therefore, we tested the repair efficiency of m1A and m3C and their byproducts with the AlkB repair enzyme. For modified structures, it is highly recommended that a biological or enzymatic assay should be adopted for identifying the product. The reason for adopting an enzymatic test is because a byproduct (e.g. m6A) may have an MW identical to the target product (e.g. m1A), which may be hard to differentiate by LC and MS analyses, including the MS/MS and exonuclease digestion with MALDI-TOF analysis.

CONCLUSION

Chemical synthesis of oligonucleotides is important for conducting various biological, chemical, and medical research including oligonucleotide drug development. The field of oligonucleotide synthesis has progressed such that DNA and RNA containing standard bases can be ordered from commercial sources. Most of these oligonucleotides are in high quality, but it is a good practice to perform quality control to confirm that the products have the correct sequences in high purity. A simple HPLC/MS analysis should be sufficient for most applications. However, a more rigorous and stringent quality control procedure should be adopted for site-specifically modified oligonucleotides, such as epigenetic marks, DNA adducts, and drug candidates. In some cases, byproducts are generated from side chemical reactions during standard automated synthesis and deprotection steps. For this reason, specific deprotection conditions may be required for preparation of oligonucleotide-containing modifications, such as epigenetic biomarkers S-hydroxymethyl-dC, S-formyl-dC, and S-carboxy-dC. After synthesis and deprotection, it is even more important to carry out a thorough purification and characterization procedure to ensure (1) complete removal of protecting groups, (2) product having high purity, and (3) modification having the correct position and identity. In this study, we used a reliable and robust procedure to characterize byproducts from m1A and m3C syntheses. The protocol used here could be helpful for identifying byproducts generated from other oligonucleotide syntheses.
**EXPERIMENTAL SECTION**

**Synthesis of Oligonucleotides Containing m1A and m3C.** Oligonucleotides (16mer) with the sequence 5′-GAAGACCTGCGCCTCC-3′ containing the lesions at the X position were made by using solid-phase phosphoramidite chemistry on a MerMade-4 Oligonucleotide synthesizer. The phosphoramidites were purchased from ChemGenes. N1-Methyl deoxyadenosine (n-fmoc) phosphoramidite was used for m1A synthesis. N3-Methyl deoxycytidine (n-bz) phosphoramidite was used for m3C synthesis. The final cleavage and deprotection step was usually carried out by treating the oligonucleotide with concentrated aqueous ammonium hydroxide (28%) at 80 °C for 3 h. The modified cleavage and deprotection step for minimizing the byproduct formation was at 25 °C for 16 h. The concentration of an oligonucleotide was determined by measuring UV absorbance at 260 nm. The extinction coefficient (ε) of a certain adduct is calculated as its unmodified counterpart because of the negligible variation between the values in the context of a 16mer DNA.

**Purity Test of Oligonucleotides Containing m1A and m3C and Related Byproducts by HPLC.** The purity of oligonucleotides was tested by both reverse-phase (C18) and anion-exchange HPLC methods. For the reverse-phase test, liquid chromatographic separation was performed by using a Phenomenex Luna Semi-Preparative C18 column (9 × 250 mm, 5 μm) at a flow rate of 3 mL/min. Solvent A was 100 mM triethylammonium acetate in water, and solvent B was 100% acetonitrile. A solvent gradient was carried out under the following conditions: 2.0% of B for 0.2 min, 2.0 to 9.0% of B over 0.3 min, 9.0 to 9.4% of B over 16 min, 9.4 to 70.0% of B over 0.5 min, 70.0% of B for 3 min, 70.0 to 2.0% of B over 0.5 min, and 2.0% B over 4.5 min. The column oven was set at 40 °C. The UV signal at 260 nm was used to detect the oligonucleotide absorbance. For the anion-exchange LC analysis, oligonucleotides were purified and tested by using a Thermo DNAPac PA-100 anion-exchange column (4 × 250 mm, 13 μm) with solvent A as water and solvent B as 1.5 M ammonium acetate in water. A solvent gradient was carried out under the following conditions: 50.0% of B for 1 min, 50.0 to 52.0% of B over 2 min, 52.0 to 75.0% over 1 min, 75.0% of B for 2 min, 75.0 to 50.0% of B over 0.5 min, and 50.0% of B for 4.5 min. The flow rate was at 4.0 mL/min.

**HPLC/MS Analysis.** HR-MS analyses of oligonucleotides were performed on AB Sciex triple quadrupole-TOF mass spectrometers (ABI4600 and ABI5600). ESI was conducted by using a needle voltage of 4.0 kV in a negative ion mode. A heated capillary was set at 600 °C. The nebulizer gas pressure was 40 psi; the heater gas pressure was 40 psi; the declustering potential was −220 V; and the collision energy was −10 V. Liquid chromatographic separation was performed using a Phenomenex Luna C18 column (4.6 × 100 mm, 5 μm) at a flow rate of 0.4 mL/min. Solvent A was 10 mM ammonium acetate in water, and solvent B was 100% acetonitrile. A solvent gradient was carried out under the following conditions: 2.0% of B for 0.5 min, 2.0 to 17.4% of B over 11 min, 17.4 to 60.0% of B over 0.1 min, 60.0% of B for 2 min, 60.0 to 2.0% of B over 0.1 min, and 2.0% B over 3.3 min.

**MS/MS Analysis.** Oligonucleotide fragmentation analyses were performed by manually injecting the oligonucleotide samples of 100 pmol/μL into AB Sciex triple quadrupole-TOF mass spectrometers (ABI4600 and ABI5600). The syringe flow rate was set at 10 μL/min. ESI was conducted by using a needle voltage of 4.5 kV in a negative ion mode. A heated capillary was set at 400 °C. The nebulizer gas pressure was 75 psi; the heater gas pressure was 25 psi; the curtain gas pressure was 25 psi; the declustering potential was −100 V; and the collision energy was −15 V. The parent ion m/z for m6A, m1A, m3C, and m3U were selected as 816.14 (−6 charge), 816.14 (−6 charge), 974.77 (−5 charge), and 974.96 (−5 charge), respectively. Data analyses were performed with the AB Sciex Analyst TF software 1.7.

**Exonuclease Digestion with MALDI-TOF Analysis.** The modified oligonucleotides were characterized by 3′−5′exonuclease digestion followed by MALDI-TOF analysis. In general, 1.0 μL of sample containing 200−250 pmol of a modified oligonucleotide was used for digestion. SVP (0.2 unit) was added together with 6.0 μL of ammonium citrate (100 mM, pH 9.4) and 6.0 μL of water for the 3′ to 5′ digestion. For the MALDI-TOF analyses, 1.0 μL of the mixture was used at a certain time point until the digestion was almost finished; the digestion was quenched by mixing with 1.0 μL of the MALDI analysis matrix (3-hydroxypicolinic acid and dianion hydrogen citrate in a 1:1 ratio). Samples were analyzed by a Shimadzu Axima Performance MALDI-TOF mass spectrometer using a 50 Hz laser with a power setting of 90 in a positive ion reflection mode with 500 shots collected.

**Enzymatic Reaction with the AlkB Protein and HPLC Analysis.** All reactions were performed at 37 °C for 1 h in a reaction buffer containing 70.0 μM Fe(NH4)2(SO4)2·6H2O, 0.93 mM α-ketoglutarate, 1.86 mM ascorbic acid, and 46.5 mM HEPES (pH 8.0). The reactions were quenched by adding 10.0 mM EDTA followed by heating at 95 °C for 5 min. Typically, 0.25 μM of the purified AlkB protein was incubated with 5.0 μM m6A and m1A in the presence of all cofactors in a 20 μL reaction volume. For m3C and m3U, 0.18 μM of AlkB was incubated with 5.0 μM of oligonucleotides. Samples were then analyzed under an HPLC condition that was able to separate the substrate and the product.

To separate the starting material and the product of the enzymatic reactions for the four lesions, different HPLC conditions were optimized by using either C18 or anion-exchange columns. The UV detection was set at 260 nm. Specifically, (1) m6A and its repaired product A were analyzed by using a Phenomenex Luna Semi-Preparative C18 column (10 × 150 mm, 5 μm) under the following conditions: solvent A was 100 mM triethylammonium acetate in water, and solvent B was 100% acetonitrile. A solvent gradient was carried out under the following conditions: 2.0% of B for 0.2 min, 2.0 to 9.0% of B over 0.3 min, 9.0 to 9.4% of B over 16 min, 9.4 to 70.0% of B over 0.5 min, 70.0% of B for 3 min, 70.0 to 2.0% of B over 0.5 min, and 2.0% B over 4.5 min. The flow rate was at 3.0 mL/min. The column oven was set at 40 °C. This protocol was also used to analyze m3U and its repair product U. (2) m1A and its repaired product A were analyzed by using a Thermo DNAPac PA-100 anion-exchange column (4 × 250 mm, 13 μm) with solvent A as water and solvent B as 1.5 M ammonium acetate in water. A solvent gradient was carried out under the following conditions: 50.0% of B for 1 min, 50.0 to 52.0% of B over 2 min, 52.0 to 75.0% over 1 min, 75.0% of B for 2 min, 75.0 to 50.0% of B over 0.5 min, and 50.0% of B for 4.5 min. The flow rate was at 4.0 mL/min. ESI was conducted by using a needle voltage of 4.0 kV in a negative ion mode. A heated capillary was set at 400 °C. The nebulizer gas pressure was 75 psi; the heater gas pressure was 25 psi; the declustering potential was −100 V; and the collision energy was −15 V. The parent ion m/z for m6A, m1A, m3C, and m3U were selected as 816.14 (−6 charge), 816.14 (−6 charge), 974.77 (−5 charge), and 974.96 (−5 charge), respectively. Data analyses were performed with the AB Sciex Analyst TF software 1.7.
over 1 min, 57.0% of B for 2 min, 57.0 to 50.0% of B for 0.5 min, and then 50.0% of B for 4.5 min.

**REFERENCES**

The authors want to thank the RI-INBRE program, its director Prof. Zahir Shaikh and the staff Kim Andrews and Patricia Murray for their kind help. We also want to thank Prof. John M. Essigmann, Prof. Bingfang Yan, Prof. David Rowley, and Kerri Bradshaw for helpful discussions.

**ACKNOWLEDGMENTS**

The authors want to thank the RI-INBRE program, its director Prof. Zahir Shaikh and the staff Kim Andrews and Patricia Murray for their kind help. We also want to thank Prof. John M. Essigmann, Prof. Bingfang Yan, Prof. David Rowley, and Kerri Bradshaw for helpful discussions.

**AUTHOR INFORMATION**

*Present Address* 
Xiamen University, 4221 Xiang An South Road, Xiang An District, Xiamen, Fujian 361102, P.R. China (F.C.).

**Author Contributions**

D.T. and A.C. contributed equally to this work.

**Funding**

This work was supported by an Institutional Development Award from the National Institute of General Medical Sciences of the National Institutes of Health under grant numbers 2 P2O GM103430, 1R15CA121042-01 (to D.L.), and a Medical Research Funds grant from the Rhode Island foundation (to D.L.). This work was in part supported by the National Institutes of Health grant CA098296 (to B.P.C.).

**Notes**

The authors declare no competing financial interest.

**AUTHOR INFORMATION**

*Corresponding Author* 
*E-mail: deyuli@uri.edu* (D.L.).

**ORCID**

Deyu Li: 0000-0002-1082-0569

**Present Address**

Xiamen University, 4221 Xiang An South Road, Xiang An District, Xiamen, Fujian 361102, P.R. China (F.C.).

**Author Contributions**

D.T. and A.C. contributed equally to this work.

**Funding**

This work was supported by an Institutional Development Award from the National Institute of General Medical Sciences of the National Institutes of Health under grant numbers 2 P2O GM103430, 1R15CA121042-01 (to D.L.), and a Medical Research Funds grant from the Rhode Island foundation (to D.L.). This work was in part supported by the National Institutes of Health grant CA098296 (to B.P.C.).

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

The authors want to thank the RI-INBRE program, its director Prof. Zahir Shaikh and the staff Kim Andrews and Patricia Murray for their kind help. We also want to thank Prof. John M. Essigmann, Prof. Bingfang Yan, Prof. David Rowley, and Kerri Bradshaw for helpful discussions.

**REFERENCES**

(1) Caruthers, M. Gene synthesis machines: DNA chemistry and its uses. *Science* 1985, 230, 281–285.

(2) Beaucage, S. L.; Caruthers, M. H. Deoxynucleoside phosphoramidites—a new class of key intermediates for deoxypolyribonucleotide synthesis. *Tetrahedron Lett.* 1981, 22, 1859–1862.

(3) Matteucci, M. D.; Caruthers, M. H. Synthesis of deoxyoligonucleotides on a polymer support. *J. Am. Chem. Soc.* 1981, 103, 3185–3191.

(4) Leumann, C. DNA analogues: from supramolecular principles to biological properties. *Bioorg. Med. Chem.* 2002, 10, 841–854.

(5) Beaucage, S. L.; Iyer, R. P. Advances in the Synthesis of Oligonucleotides by the Phosphoramidite Approach. *Tetrahedron* 1992, 48, 2223–2311.

(6) Petersen, M.; Wengel, J. LNA: a versatile tool for therapeutics and genomics. *Trends Biotechnol.* 2003, 21, 74–81.

(7) De Mesmaeker, A.; Altman, K.-H.; Waldner, A.; Wendeborn, S. Backbone modifications in oligonucleotides and peptide nucleic acid systems. *Curr. Opin. Struct. Biol.* 1995, 5, 343–355.

(8) Antisense Drug Technology: Principles, Strategies, and Applications, 2nd ed.; Crooke, S. T., Ed.; CRC Press: Boca Raton, 2008.

(9) *Handbook of Analysis of Oligonucleotides and Related Products*, 1st ed; Bonilla, J. V., Ed.; CRC Press: Boca Raton, 2016.

(10) Chen, F.; Tang, Q.; Bian, K.; Humulock, Z. T.; Yang, X.; Jost, M.; Drennan, C. L.; Essigmann, J. M.; Li, D. Adaptive response enzyme AlkB preferentially repairs 1-methylguanine and 3-methylthymine adducts in double-stranded DNA. *Chem. Res. Toxicol.* 2016, 29, 687–693.

(11) Jain, V.; Hilton, B.; Lin, B.; Jain, A.; MacKerell, A. D.; Zou, Y.; Cho, B. P. Structural and thermodynamic insight into *Escherichia coli* UvrABC-mediated incision of cluster diacylaminoazofluorene adducts on the NarI sequence. *Chem. Res. Toxicol.* 2013, 26, 1251–1262.

(12) Jain, V.; Vaidyanathan, V. G.; Patnaik, S.; Gopal, S.; Cho, B. P. Conformational insights into the lesion and sequence effects for alylamine-induced translesion DNA synthesis: 1H NMR, surface plasmon resonance, and primer kinetic studies. *Biochemistry* 2014, 53, 4059–4071.

(13) Sandineni, A.; Lin, B.; MacKerell, A. D.; Cho, B. P. Structure and thermodynamic insights on acetylaminoazofluorene-modified deletion DNA duplexes as models for frameshift mutagenesis. *Chem. Res. Toxicol.* 2013, 26, 957–951.

(14) Jain, N.; Li, Y.; Zhang, L.; Meneni, S. R.; Cho, B. P. Probing the sequence effects on NarI-induced 2 frameshift mutagenesis by dynamic 19F NMR, UV, and CD spectroscopy. *Biochemistry* 2007, 46, 13310–13321.

(15) *Molecular Biology of Mutagens and Carcinogens*, 1st ed.; Singer, B., Ed.; Springer US: Boston, MA, 1983.

(16) Beranek, D. T.; Weis, C. C.; Swenson, D. H. A comprehensive quantitative analysis of methylated and ethylated DNA using high pressure liquid chromatography. *Carcinogenesis* 1980, 1, 595–606.

(17) Boiteux, S.; Laval, J. Mutagenesis by alkylating agents: coding properties for DNA polymerase of poly (dC) template containing 3-methylcytosine. *Biochimie* 1982, 64, 637–641.

(18) Gomes, J. C.; Chang, C.-J. Reverse-phase high-performance liquid chromatography of chemically modified DNA. *Anal. Biochem.* 1983, 129, 387–391.

(19) Chang, C. J.; Gomes, J. D.; Byrn, S. R. Chemical modification of deoxyribonucleic acids: a direct study by carbon-13 nuclear magnetic resonance spectroscopy. *J. Org. Chem.* 1983, 48, 5151–5160.

(20) Kawasaki, H.; Ninomiya, S.; Yuki, H. High-performance liquid chromatographic determination of 3-methylcytosine in deoxyribonucleic acid treated with carcinogenic methylating agents in vitro and in vivo. *Chem. Pharm. Bull.* 1985, 33, 1170–1174.

(21) Ashworth, D. J.; Baird, W. M.; Chang, C.-J.; Ciupek, J. D.; Busch, K. L.; Cooks, R. G. Chemical modification of nucleic acids. Methylation of calf thymus DNA investigated by mass spectrometry and liquid chromatography. *Biomed. Mass Spectrom.* 1985, 12, 309–318.

(22) Kelley, M. R.; Fishel, M. L. DNA repair proteins as molecular targets for cancer therapeutics. *Anticancer Agents Med. Chem.* 2008, 8, 417–425.

(23) Messouidi, K.; Clavreul, A.; Lagarce, F. Toward an effective strategy in globlastoma treatment. Part I: resistance mechanisms and strategies to overcome resistance of globlastoma to temozolomide. *Drug Discovery Today* 2015, 20, 899–905.

(24) Culp, L. A.; Dore, E.; Brown, G. M. Methylated bases in DNA of animal origin. *Arch. Biochem. Biophys.* 1970, 136, 73–79.

(25) Margison, G. P.; Margison, J. M.; Montesano, R. Methylated purines in the deoxyribonucleic acid of various Syrian-golden-hamster tissues after administration of a hepatocarcinogenic dose of dimethylnitrosamine. *Biochem. J.* 1976, 157, 627–634.

(26) Frei, J. V.; Swenson, D. H.; Warren, W.; Lawley, P. D. Alkylation of deoxyribonucleic acid in vivo in various organs of CS7BL mice by the carcinogens N-methyl-N-nitrosourea, N-ethyl-N-nitrosourea and ethyl methanesulphonate in relation to induction of thymic lymphoma. Some applications of high-pressure liquid chromatography. *Biochem. J.* 1978, 174, 1031–1044.

(27) Faustman, E. M.; Goodman, J. L. A method for the rapid quantitation of methylated hepatic DNA-purines using high pressure liquid chromatography. *J. Pharmacol. Methods* 1990, 4, 303–312.

(28) Beranek, D. T.; Heffich, R. H.; Kodel, R. L.; Morris, S. M.; Casciano, D. A. Correlation between specific DNA-methylation products and mutation induction at the HGPRT locus in Chinese hamster ovary cells. *Mutat. Res.* 1983, 110, 171–180.
(29) Faustman-Watts, E. M.; Goodman, J. I. DNA-purine methylation in hepatic chromatin following exposure to dimethylnitrosamine or methylmethylnitrosourea. *Biochem. Pharmacol.* 1984, 33, 585–590.
(30) Trewick, S. C.; Henshaw, T. F.; Hausinger, R. P.; Lindahl, T.; Sedgwick, B. Oxidative demethylation by *Escherichia coli* AlkB directly reverses DNA base damage. *Nature* 2002, 419, 174–178.
(31) Falnes, P. O.; Johansen, R. F.; Seeborg, E. AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*. *Nature* 2002, 419, 178–182.
(32) Fedele, B. I.; Singh, V.; Delaney, J. C.; Li, D.; Essigmann, J. M. The AlkB family of Fe(II)/α-ketoglutarate-dependent dioxygenases: repairing nucleic acid alkylation damage and beyond. *J. Biol. Chem.* 2015, 290, 20734–20742.
(33) Sedgwick, B.; Bates, P.; Paik, J.; Jacobs, S.; Lindahl, T. Repair of alkylated DNA: recent advances. *DNA Repair* 2007, 6, 429–442.
(34) Delaney, J. C.; Essigmann, J. M. Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 2004, 101, 14051–14056.
(35) Patalot-Meakin, T.; Pillay, N.; Beck, S. 3-methylthymine in cancer: an underappreciated methyl lesion? Epigenomics 2016, 8, 451–454.
(36) Delaney, J. C.; Essigmann, J. M. Biological properties of single chemical–DNA adducts: a twenty year perspective. *Chem. Rev. Toxicol.* 2008, 21, 232–252.
(37) Zeng, Y.; Wang, Y. Facile Formation of an Intrastrand Cross-Link Lesion between Cytosine and Guanine upon Pyrex-Filtered UV Light Irradiation of d(5CG) and Duplex DNA Containing S-Bromocytosine. *J. Am. Chem. Soc.* 2004, 126, 6552–6553.
(38) Gao, H.; Jiang, Y.; Wang, Y. Stereospecific Synthesis and Characterization of Oligodeoxiribonucleotides Containing an N8-(1-Carboxyethyl)-2′-deoxyguanosine. *J. Am. Chem. Soc.* 2007, 129, 12123–12130.
(39) Li, D.; Delaney, J. C.; Page, C. M.; Chen, A. S.; Wong, C.; Drennan, C. L.; Essigmann, J. M. Repair of DNA alkylation damage by the *Escherichia coli* adaptive response protein AlkB as studied by ESI-TOF mass spectrometry. *J. Nucleic Acids* 2010, 10, 369434.
(40) Gao, L.; Zhang, L.; Cho, B. P.; Chiarelli, M. P. Sequence verification of oligonucleotides containing multiple arylamine modifications by enzymatic digestion and liquid chromatography mass spectrometry (LC/MS). *J. Am. Soc. Mass Spectrom.* 2008, 19, 1147–1155.
(41) Zhang, L.-K.; Gross, M. L. Matrix-assisted laser desorption/ionization mass spectrometry methods for oligodeoxynucleotides: improvements in matrix, detection limits, quantification, and sequencing. *J. Am. Soc. Mass Spectrom.* 2000, 11, 854–866.
(42) Redon, S.; Bombard, S.; Elizondo-Riojas, M.-A.; Chottard, J.-C. Platinum cross-linking of adenines and guanines on the quadruplex structures of the AG3(T2AG3)3 and (T2AG3)4 human telomere sequences in Na+ and K+ solutions. *Nucleic Acids Res.* 2003, 31, 1605–1613.
(43) Wang, B.; Sayer, J. M.; Yagi, H.; Frank, H.; Seidel, A.; Jerina, D. M. Facile interstrand migration of the hydrocarbon moiety of a dibenzo[a,l]pyrene 11,12-diol 13,14-epoxide adduct at N2 of deoxyguanosine in a duplex oligonucleotide. *J. Am. Chem. Soc.* 2006, 128, 10079–10084.
(44) Tretjakova, N.; Matter, B.; Ogdie, A.; Wishnok, J. S.; Tannenbaum, S. R. Locating nucleobase lesions within DNA sequences by MALDI-TOF mass spectral analysis of exonuclease ladders. *Chem. Rev. Toxicol.* 2001, 14, 1058–1070.
(45) Engel, J. D. Mechanism of the Dimroth rearrangement in adenosine. *Biochem. Biophys. Res. Commun.* 1975, 64, 581–586.
(46) Leutzinger, E. E.; Miller, P. S.; Kan, L.-S. Studies on the hydrolysis of 3-methyl-2′-deoxycytidine in aqueous solution A synthesis of 3-methyl-2′-deoxouridine. *Biochim. Biophys. Acta, Gene Struct. Expression* 1982, 697, 243–251.
(47) Frick, L. E. The versatile *E. coli* adaptive response protein AlkB mitigates toxicity and mutagenicity of etheno-, ethano-, and methyl-modified bases in vivo. Ph.D. Thesis, Massachusetts Institute of Technology, 2007.
(48) Li, D.; Fedele, B. I.; Shrivastav, N.; Delaney, J. C.; Yang, X.; Wong, C.; Drennan, C. L.; Essigmann, J. M. Removal of N-alkyl modifications from N²-alkylguanine and N³-alkylcytosine in DNA by the adaptive response protein AlkB. *Chem. Res. Toxicol.* 2013, 26, 1182–1187.
(49) Li, D.; Delaney, J. C.; Page, C. M.; Yang, X.; Chen, A. S.; Wong, C.; Drennan, C. L.; Essigmann, J. M. Exocyclic carbons adjacent to the N² of adenine are targets for oxidation by the *Escherichia coli* adaptive response protein AlkB. *J. Am. Chem. Soc.* 2012, 134, 8896–8901.
(50) Singh, V.; Fedele, B. I.; Li, D.; Delaney, J. C.; Kozekov, I. D.; Kozekova, A.; Marnett, L. J.; Rizzo, C. J.; Essigmann, J. M. Mechanism of repair of acrolein- and malondialdehyde-derived exocyclic guanine adducts by the α-ketoglutarate/Fe(II) dioxygenase AlkB. *Chem. Res. Toxicol.* 2014, 27, 1619–1631.
(51) Basiri, B.; Bartlett, M. G. LC−MS of oligonucleotides: applications in biomedical research. *Bioanalysis* 2014, 6, 1525–1542.
(52) van Dongen, W. D.; Niessen, W. M. A. Bioanalytical LC-MS of therapeutic oligonucleotides. *Bioanalysis* 2011, 3, 541–564.
(53) Lin, Z. J.; Li, W.; Dai, G. Application of LC-MS for quantitative analysis and metabolite identification of therapeutic oligonucleotides. *J. Pharm. Biomed. Anal.* 2007, 44, 330–341.
(54) Schröder, A. S.; Steibacher, J.; Steigenberger, B.; Gnerlich, F. A.; Schiesser, S.; Pfaffeneder, T.; Carell, T. Synthesis of a DNA promoter segment containing all four epigenetic nucleosides: 5-methyl-, 5-hydroxymethyl-, 5-formyl-, and 5-carboxy-2′-deoxycytidine. *Angew. Chem., Int. Ed. Engl.* 2014, 53, 315–318.
(55) Münzel, M.; Globisch, D.; Trindler, C.; Carell, T. Efficient synthesis of 5-hydroxymethylcytosine containing DNA. *Org. Lett.* 2010, 12, 5671–5673.
(56) Delaney, J. C.; Smeester, L.; Wong, C.; Frick, L. E.; Taghizadeh, K.; Wishnok, J. S.; Drennan, C. L.; Samson, L. D.; Essigmann, J. M. AlkB reverses etheno DNA lesions caused by lipid oxidation in vitro and in vivo. *Nat. Struct. Mol. Biol.* 2005, 12, 855–860.